Topological Data Analysis Spring 2020
Magnus Bakke Botnan

Contents

1 Week 1: Introduction and Simplicial Complexes ........................................... 3
1.1 Simplices .......................................................................................... 3
1.2 Simplicial Complexes .......................................................................... 4
1.3 Abstract Simplicial Complexes ............................................................ 5
1.4 Triangulations ..................................................................................... 5
1.5 The Čech and Vietoris–Rips Simplicial Complexes .................................. 6

2 Week 2: Simplicial Homology ................................................................. 7
2.1 Free Vector Spaces ............................................................................ 9
2.2 The Betti Number and the Euler Characteristic .................................. 10
2.3 A Remark on the Choice of Coefficients ........................................... 12
2.4 Homology .......................................................................................... 13

3 Week 3: Computation, Induced Maps and Invariance ............................. 15
3.1 The Standard Algorithm .................................................................. 15
3.1.1 The Homology of a Triangle ....................................................... 15
3.1.2 The Algorithm ............................................................................ 16
3.2 Simplicial Maps and their Induced Maps in Homology ....................... 17
3.3 Simplicial Approximations and Invariance of Homology .................... 18
3.3.1 Simplicial Maps Revisited ......................................................... 18
3.3.2 Simplicial Approximations ....................................................... 18

4 Week 4: Homotopy Equivalence, Relative Homology and Sensor Networks 21
4.1 Homotopy Equivalence .................................................................... 21
4.2 The Nerve Theorem ......................................................................... 23
4.2.1 The Čech Complex .................................................................... 24
4.3 Relative Homology .......................................................................... 24
4.4 Sensor Networks ............................................................................. 26
4.4.1 The covering problem ............................................................... 26

5 Week 5: Topology & Homology Inference and Persistent Betti Numbers 28
5.1 Topology & Homology Inference ...................................................... 28
5.2 Persistent Betti Numbers .................................................................. 29
5.2.1 The Čech and Vietoris–Rips Filtrations ....................................... 32

6 Week 6: Algebraic Foundations ............................................................ 34
6.1 Motivation ......................................................................................... 34
6.2 Persistence Modules ......................................................................... 35
<table>
<thead>
<tr>
<th>Week 7: The Persistence Algorithm</th>
<th>40</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.1 The standard algorithm</td>
<td>40</td>
</tr>
<tr>
<td>7.2 Clearing</td>
<td>43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week 9: Cohomology</th>
<th>45</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1 Dual persistence modules</td>
<td>45</td>
</tr>
<tr>
<td>8.2 Simplicial Cohomology</td>
<td>46</td>
</tr>
<tr>
<td>8.3 Persistent Cohomology</td>
<td>48</td>
</tr>
<tr>
<td>8.3.1 Computation</td>
<td>49</td>
</tr>
<tr>
<td>8.3.2 Clearing</td>
<td>50</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week 10: Stability of Persistent Homology</th>
<th>52</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.1 The Bottleneck Distance</td>
<td>52</td>
</tr>
<tr>
<td>9.1.1 Stability of Persistent Homology</td>
<td>53</td>
</tr>
<tr>
<td>9.1.2 Generalizations</td>
<td>55</td>
</tr>
<tr>
<td>9.2 Interleaveings</td>
<td>55</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week 11: Clustering and TDA</th>
<th>59</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.1 Some Examples and Kleinberg’s Theorem</td>
<td>59</td>
</tr>
<tr>
<td>10.2 ToMATo</td>
<td>61</td>
</tr>
<tr>
<td>10.2.1 The Discrete Case</td>
<td>62</td>
</tr>
<tr>
<td>10.3 A First Taste of Mapper</td>
<td>65</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week 12: Reeb Graphs and Mapper Continued</th>
<th>66</th>
</tr>
</thead>
<tbody>
<tr>
<td>11.1 Reeb Graphs</td>
<td>66</td>
</tr>
<tr>
<td>11.2 The Continuous Mapper</td>
<td>66</td>
</tr>
<tr>
<td>11.3 Mapper in Practice</td>
<td>68</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week 13: Generalized Persistence</th>
<th>71</th>
</tr>
</thead>
<tbody>
<tr>
<td>12.1 Zigzag Persistent Homology</td>
<td>71</td>
</tr>
<tr>
<td>12.1.1 Levelset Zigzag Persistent Homology</td>
<td>72</td>
</tr>
<tr>
<td>12.1.2 Reeb Graphs</td>
<td>73</td>
</tr>
<tr>
<td>12.2 Multiparameter Persistent Homology</td>
<td>74</td>
</tr>
<tr>
<td>12.2.1 RIVET</td>
<td>77</td>
</tr>
<tr>
<td>12.2.2 Distances</td>
<td>78</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Week 14: $d_I = d_B$</th>
<th>81</th>
</tr>
</thead>
<tbody>
<tr>
<td>13.1 $d_I &lt; d_B$</td>
<td>81</td>
</tr>
<tr>
<td>13.2 $d_B &lt; d_I$</td>
<td>82</td>
</tr>
<tr>
<td>13.2.1 Hall’s Theorem</td>
<td>85</td>
</tr>
<tr>
<td>13.2.2 Wrapping Up</td>
<td>86</td>
</tr>
</tbody>
</table>
1 Week 1: Introduction and Simplicial Complexes

The idea of Topological Data Analysis (TDA) is to assign topological invariants to data. Data however is typically presented as a discrete sample whose topology is rather uninteresting. Therefore we must transform the data into a “continuous” object which (hopefully) is topologically similar to the shape of the underlying geometric object from which the data is sampled.

Take for example the discrete set of data points in the figure below. Clearly there is ”circular structure” to the data and our pattern-seeking brains have no trouble inferring such information. Topology per-se does not measure this circularity but after thickening every point by a certain predefined thickness an annulus appears. This is a topological object with a well-defined ”circular structure”.

In general, the resulting continuous spaces need to be discretized before we can feed them into our computers, and this is typically done by means of simplicial complexes. In the first lecture we discuss simplicial complexes in general and two constructions in particular: the Čech and Vietoris–Rips complexes.

Most of the following material can be found in [10, Sections 1.1-3] and [8, III.1-3].

1.1 Simplices

A set of points \( \{p_0, \ldots, p_n\} \subset \mathbb{R}^d \) is said to be geometrically independent\(^1\) if for any (real) scalars \( \{t_i\}_{i=0}^n \), the equations

\[
\sum_{i=0}^n t_i = 0 \quad \quad \sum_{i=0}^n t_i a_i = 0
\]

imply that \( t_0 = t_1 = \ldots = t_n = 0 \). This is clearly equivalent to the vectors \( \{p_1 - p_0, \ldots, p_n - p_0\} \) being linearly independent.

**Definition 1.1.** Let \( \{p_0, \ldots, p_n\} \) be a geometrically independent set in \( \mathbb{R}^d \). We define the \( n \)-simplex spanned by the points \( p_i \) to be the set of all points \( x \in \mathbb{R}^d \) of the form

\[
x = \sum_{i=0}^n t_i p_i
\]

where \( \sum_{i=0}^n t_i = 1 \), and \( t_i \geq 0 \) for all \( i \). The point \( x \) is a **convex combination** of the points \( p_0, \ldots, p_n \), and the uniquely determined numbers \( t_i \) are called the **barycentric coordinates**.

One immediately observes that any point is a 0-simplex, any two distinct points form a 1-simplex (a line), three non-collinear points form a 2-simplex (a triangle), four non-coplaner points form a 3-simplex (a tetrahedron) and so forth. Note that there can be no \( n \)-simplex in \( \mathbb{R}^d \) for \( n > d \). The points \( p_0, \ldots, p_n \) that span a simplex \( \sigma \) are called the **vertices** of \( \sigma \); the number

\(^1\)Or affinely independent.
n is the dimension of $\sigma$. A simplex is regular if all its edges have the same length. The standard $n$-simplex $\Delta^n$ spanned by the endpoints of the unit vectors along the coordinate axes in $\mathbb{R}^{n+1}$ is a regular simplex.

By definition, any subset of $\{p_0, \ldots, p_n\}$ is again geometrically independent and thus defines a simplex in its own right. We say that $\tau$ is a face of $\sigma$ if $\tau$ is spanned by a subset of the vertices defining $\sigma$. We denote this by $\tau \subseteq \sigma$ or $\tau \subset \sigma$ if $\tau$ is a proper face of $\sigma$, i.e. $\tau \neq \sigma$. The union of the proper faces of $\sigma$ is called the boundary of $\sigma$ and is denoted by $\text{Bd}(\sigma)$. The simplex $\sigma$ is a (proper) coface of $\tau$ if $\tau$ is a (proper) face of $\sigma$.

1.2 Simplicial Complexes

By gluing simplices along faces we obtain topological spaces whose properties are determined by their combinatorial properties.

**Definition 1.2.** A simplicial complex $K$ in $\mathbb{R}^d$ is a finite collection of simplices in $\mathbb{R}^d$ such that:

1. Every face of a simplex of $K$ is in $K$.
2. The non-empty intersection of any two simplices of $K$ is a face of each of them.

The dimension of $K$ is the maximum dimension of its simplices and is denoted by $\text{dim } K$. If $L$ is a subcollection of simplices that contains all faces of its elements, then $L$ is a simplicial complex in its own right; it is called a subcomplex of $K$. The $p$-skeleton $K^{(p)}$ of $K$ is the collection of all simplices of $K$ of dimension at most $p$. In particular, $K^{(0)}$ is the vertex set of $K$.

---

2We will only consider finite simplicial complexes in this course.
Let $|K|$ denote the subset of $\mathbb{R}^d$ given by the union of all the simplices in $K$, equipped with the subspace topology. This space is called the **underlying space** of $K$. A **polyhedron** is a subset of $\mathbb{R}^d$ that can be realized as the underlying space of a simplicial complex.

### 1.3 Abstract Simplicial Complexes

Often the precise geometry of the simplicial complex is not important. We now introduce a purely combinatorial description of simplicial complexes under which two objects are isomorphic if and only if their geometric counterparts are linearly isomorphic.

**Definition 1.3.** An **abstract simplicial complex** is a finite collection $A$ of finite non-empty sets, such that if $\alpha$ is an element of $A$, then so is every nonempty subset of $\alpha$.

We call the elements of $A$ **simplices** and the **dimension** of $\alpha \in A$ is $\dim \alpha = |\alpha| - 1$. As above the dimension of $A$ is given by $\dim A = \max_{\alpha \in A} \dim \alpha$. We say that $B \subseteq A$ is a **subcomplex** of $A$ if $B$ is an abstract simplicial complex. The **vertices** $V_A$ of $A$ is the union of the one-point elements of $A$. We say that two simplical complexes $A$ and $B$ are **isomorphic** if there is a bijective correspondence $f: V_A \rightarrow V_B$ satisfying $\{a_0, \ldots, a_n\} \in A$ if and only if $\{f(a_0), \ldots, f(a_n)\} \in B$. A (geometric) simplicial complex $K$ defines an abstract simplicial complex $A$ by identifying any $K$ with its vertices. We say that $A$ is the **vertex scheme** of $K$. The line connecting $p_0$ and $p_1$ has vertex scheme $\{\{p_0\}, \{p_1\}, \{p_0, p_1\}\}$. If $A$ is the vertex scheme of $K$, then we say that $K$ is the **geometric realization** of $A$. It is easy to see that any abstract simplicial complex on $d + 1$ vertices can be geometrically realized as subcomplex of any $d$-simplex in $\mathbb{R}^d$. The next lemma shows that we often can do considerably better.

A finite subset $P \subset \mathbb{R}^d$ is in **general position** if any subset of at most $d + 1$ points is geometrically independent.

**Remark 1.4.** Most point-sets are in general position, and those who are not can be brought into general position by an arbitrary small perturbation. One may verify that any subset of at most $r + 1$ points in the image of the curve $\gamma: \mathbb{R} \rightarrow \mathbb{R}^r$ defined by $\gamma(t) = (t, t^2, \ldots, t^r)$ is in general position.

**Lemma 1.5.** An abstract simplicial complex $A$ with $\dim A = d$ has a geometric realization in $\mathbb{R}^{2d + 1}$.

**Proof.** Let $h: V_A \rightarrow \mathbb{R}^{2d + 1}$ be any function such that $h(V_A)$ is in general position. Any simplex $\alpha \in A$ has at most $d + 1$ vertices, and thus its vertices correspond to a geometrically independent set under the function $h$. We denote this set by $h(\alpha)$, and the associated geometric simplex by $\sigma_\alpha$. It remains to show that the intersection of any two simplices is a face of both the simplices.

Let $\alpha$ and $\beta$ be any two simplices and observe that $|\alpha \cup \beta| \leq 2(d + 1)$, and therefore $h(\alpha) \cup h(\beta)$ forms a geometrically independent set. It follows that any point $x \in \sigma_\alpha \cap \sigma_\beta$ can be uniquely represented as a convex combination of the vertices $h(\alpha) \cup h(\beta)$. Hence the barycentric coordinates of $x$ can be non-zero only for $t_i \in h(\alpha \cap \beta)$. We conclude that $x \in \sigma_{\alpha \cap \beta}$. \qed

### 1.4 Triangulations

**Definition 1.6.** A topological space $X$ is **triangulable** if there exists a simplicial complex $K$ whose underlying space is homeomorphic to $X$. Any homeomorphism $h: |K| \rightarrow X$ is a **triangulation**.

**Example 1.7.** Abstract simplicial complexes can be very useful in defining triangulations of spaces; see Fig. 3 for an example of how to specify a triangulation of the cylinder (of finite height).
Remark 1.8. It is a well-known fact that any closed surface can be triangulated and the same holds true for 3-dimensional manifolds. The situation becomes significantly more complicated in dimension 4 where manifolds need not admit a triangulation. The comb space is an example of a compact, path-connected subset of the plane that cannot be triangulated (why?).

1.5 The Čech and Vietoris–Rips Simplicial Complexes

We now turn our attention to two of the most widely used simplicial complexes constructed in topological data analysis. We will return to them for a more formal treatment later in the course. In what follows $P = \{p_0, \ldots, p_n\} \subset \mathbb{R}^d$.

The Čech complex. For a point $x \in \mathbb{R}^d$, let $B_r(x) := \{x' \in \mathbb{R}^d : ||x - x'|| \leq r\}$ denote the closed ball of radius $r$ centred at $x$. The Čech complex of $P$ at scale $r$ is the simplicial complex

$$\text{Čech}_r(P) = \{\sigma \subseteq P : \bigcap_{p \in \sigma} B_r(p) \neq \emptyset\}.$$ 

Remark 1.9. We shall see that $|\text{Čech}_r(P)|$ faithfully captures the topology of $\bigcup_{p \in P} B_r(P)$.

Determining if $\sigma \in \text{Čech}_r(P)$ may appear unwieldy for simplices of large cardinality. The following theorem shows that we need not consider all possible intersections as long as the embedding dimension is small.

Theorem 1.10 (Helly’s Theorem). Let $F$ denote a finite collection of closed, convex sets in $\mathbb{R}^d$. If the intersection of any $d + 1$ sets is non-empty, then the whole collection has non-empty intersection.

Proof. See e.g. [8, III.2].

Hence, if $P \subset \mathbb{R}^3$, then computing all quadruple intersections will be sufficient for determining $\text{Čech}_r(P)$. Furthermore, one observes that $\sigma \in \text{Čech}_r(P)$ if and only if there exists a ball of radius $r$ containing all the elements of $\sigma$. The unique (exercise) smallest enclosing ball containing a set of points $\sigma \subseteq P$ is called the minimal enclosing ball, and it can be effectively computed with the miniball algorithm [8, III.2].
The Vietoris–Rips complex. Recall that the diameter of a subset \( \sigma \subseteq P \) is given by \( \text{diam} \sigma = \max_{p_i, p_j \in \sigma} | |p_i - p_j|| \). The Vietoris–Rips complex of \( P \) at scale \( r \) is the simplicial complex

\[
\text{VR}_r(P) = \{ \sigma \subseteq P : \text{diam} \sigma \leq 2r \}.
\]

It is straightforward to verify that \( \text{Cech}_r(P) \subseteq \text{VR}_r(P) \subseteq \text{Cech}_{2r}(P) \). The latter bound can however be significantly sharpened.

**Theorem 1.11** (Jung’s Theorem). Let \( Q \) be a finite point set in \( \mathbb{R}^d \) with diameter \( \text{diam} Q \). Then \( Q \) is contained in a closed ball with radius \( r \leq \delta \text{diam} Q \) where \( \delta = \sqrt{\frac{d}{2(d+1)}} \).

**Proof.** Assume that the minimal enclosing ball of \( Q \) has center \( z \) and radius \( r \), and let \( S_r(z) \) denote its boundary. By minimality, \( S_r(z) \cap Q \neq \emptyset \) and we let \( \{q_1, \ldots, q_n\} \) be a maximally geometrically independent subset of \( S_r(z) \cap Q \). Note that \( n \leq d + 1 \) since the points lie in \( \mathbb{R}^d \). If \( n < d + 1 \), then extend the set with elements \( \{q_{n+1}, \ldots, q_{d+1}\} \subset S_r(z) \setminus S_r(z) \cap Q \) such that \( \{q_1, \ldots, q_{d+1}\} \) is geometrically independent. Then \( S_r(z) \) is the unique sphere passing through the \( q_i \)'s, and \( z \) can be expressed as a convex combination as follows

\[
z = \sum_{i=1}^{d+1} t_i q_i \quad \text{and} \quad \sum_{i=1}^{d+1} t_i = 1.
\]

Without loss of generality we will assume that \( z = 0 \). For a fixed \( 1 \leq k \leq d + 1 \) we get

\[
\sum_{i=1, i \neq k}^{d+1} t_i ||q_i - q_k||^2 = \sum_{i=1}^{d+1} t_i ||q_i - q_k||^2 = \sum_{i=1}^{d+1} t_i ||q_i||^2 - 2q_k \cdot \sum_{i=1}^{d+1} t_i q_i + ||q_k||^2 \sum_{i=1}^{d+1} t_i = r^2 - 0 + r^2 = 2r^2.
\]

Summing over all \( 1 \leq k \leq (d + 1) \) gives

\[
(d + 1)(2r^2) = \sum_{k=1}^{d+1} \sum_{i=1, i \neq k}^{d+1} t_i ||q_i - q_k||^2 \leq \sum_{k=1}^{d+1} \sum_{i=1, i \neq k}^{d+1} t_i (\text{diam} Q)^2 = \sum_{k=1}^{d+1} (1 - t_k)(\text{diam} Q)^2 = d(\text{diam} Q)^2.
\]

We conclude that

\[
r \leq \sqrt{\frac{d}{2(d+1)}} \cdot \text{diam} Q.
\]

The bound is tight for regular \( d \)-simplices in \( \mathbb{R}^d \) (exercise).

**Corollary 1.12.** \( \text{Cech}_r(P) \subseteq \text{VR}_r(P) \subseteq \text{Cech}_{2\delta r}(P) \) where \( 2\delta = 2\sqrt{\frac{d}{2(d+1)}} = \sqrt{\frac{2d}{d+1}} \).

## 2 Week 2: Simplicial Homology

Distinguishing topological spaces is a notoriously difficult problem: deciding if two simplicial complexes are homeomorphic (or homotopy equivalent) is an **undecidable problem**. Roughly speaking, this means that there exists no algorithm which takes as input two simplicial complexes and outputs ”true” if and only if they are homeomorphic. Hence, to show that two
Figure 4: The Euler characteristic counts the number of circular holes.

spaces are not homeomorphic, it is often more fruitful to show that they look different under the lens of a tool which is insensitive to homeomorphism (or other forms of equivalence). Algebraic topology offers one such approach: algebraic invariants such as homology and homotopy groups are associated to topological spaces. We shall focus on the former of these two as its computation reduces to simple linear algebra whereas the latter is exceptionally hard to compute. Although the outputted homology invariants can be hard to interpret, we shall see that for low-dimensional simplicial complexes, the invariants have a clear interpretation in terms of connectivity, "circularity" and cavities. Our first topological invariant, the Euler characteristic, will illustrate this.

Definition 2.1. Let $K$ be a simplicial complex and let $K_i$ denote the number of $i$-simplices in $K$. The Euler characteristic of $K$ is the integer

$$\chi(K) = \sum_{i \geq 0} (-1)^i K_i.$$

It is not at all clear that this number depends only on the homeomorphism type of the underlying space $|K|$, and we will return to that fact in the next lecture. For now, we will try to see what it represents by considering some familiar topological spaces.

Example 2.2. Consider the planar simplicial complex of Fig. 4. Its Euler characteristic is

$$\chi = \text{#vertices} - \text{#edges} + \text{#triangles} = 11 - 17 + 3 = -3.$$

We see that $|\chi|$ is one less than the number of "holes" in the graph. The reader may verify that the same relation holds for other planar simplicial complexes. This is no coincidence: in the exercises we saw that for a connected planar graph the following is always true

$$2 = \text{#vertices} - \text{#edges} + \text{#faces}.$$

The number of faces may in turn be added up as follows:

$$\text{#faces} = \text{#triangles} + \text{#holes} + 1(\text{the unbounded component}).$$

Hence,

$$1 - \text{#holes} = \text{#vertices} - \text{#edges} + \text{#triangles} = \chi.$$

Example 2.3. The regular tetrahedron and regular icosahedron are two of the five platonic solids. They are both triangulations of the sphere and one can easily verify that their Euler characteristics coincide:

$$\chi(\text{tetrahedron}) = 4 - 6 + 4 = 2$$

$$\chi(\text{icosahedron}) = 12 - 30 + 20 = 2.$$
The torus on the other hand has a triangulation as shown in Fig. 5. The Euler characteristic of the triangulation is \( \chi = 9 - 27 + 18 = 0 \). Again this is no coincidence: for a simplicial complex \( K \) embedded in \( \mathbb{R}^3 \) the following relation holds:

\[
\chi(K) = \text{#connected components} - \text{#independent loops} + \text{#voids}.
\]

There are no non-trivial loops on the sphere, whereas the torus has two, as depicted in Fig. 5.

The fact that the Euler Characteristic is invariant under homeomorphism allows us one talk about the Euler characteristic of a triangulated topological space. Hence, the previous example shows that the torus and the sphere are not homeomorphic.

**Remark 2.4.** Remarkably, two closed, oriented surfaces are homeomorphic if and only if they have the same Euler characteristic: the Euler characteristic of a genus \( g \) surface is \( 2 - 2g \).

In light of the introduction above, this can clearly not be the case for manifolds of arbitrary dimension (why?).

### 2.1 Free Vector Spaces

We briefly recap the construction of a free vector space on a finite set. Let \( S \) be a finite set and \( k \) an arbitrary field. We define the **free \( k \)-vector space generated by \( S \)** to be the vector space \( F(S) \) with elements given by formal linear combinations \( \sum_{s_i \in S} a_i s_i \) where each \( a_i \in k \) and \( s_i \in S \). Here addition is defined component-wise in the obvious way:

\[
\sum_{s_i \in S} a_i s_i + \sum_{s_i \in S} b_i s_i = \sum_{s_i \in S} (a_i + b_i) s_i,
\]

where \( a_i + b_i \in k \). The resulting \( k \)-vector space has a canonical vector space basis given by the elements in \( S \).

**Example 2.5.** Let \( k = \mathbb{Z}_2 \) be the field of two elements and \( S = \{a, b, c\} \) a set of three elements. Then \( F(S) \) is the 3-dimensional vector space with elements \( \{0, a, b, c, a + b, a + c, b + c\} \) and a basis \( \{a, b, c\} \). Addition is straight-forward: \( (a + b) + (a + c) = (2a) + b + c = b + c \).

**Example 2.6.** If we instead let \( k = \mathbb{Z}_3 \) and \( S = \{a, b\} \) then

\[
F(S) = \{0, a, 2a, b, 2b, a + b, a + 2b, 2a + b, 2a + 2b\}.
\]

We observe that \( F(S) \) has dimension \( |S| \) and cardinality \( |S|^{|k|} \).
2.2 The Betti Number and the Euler Characteristic

We shall now construct a family of topological invariants which measure “connectedness” in a given dimension. By means of this we can make the interpretation of the Euler characteristic precise. For the time being we shall assume that $k = \mathbb{Z}_2$.

**Definition 2.7.** Let $n \geq 0$ be an integer and $K$ a simplicial complex. The vector space of $n$-chains in $K$ is the free $\mathbb{Z}_2$-vector space $C_n(K)$ generated by the $n$-simplices of $K$.

By defining the boundary of an $n$-simplex to be the formal sum of its $(n-1)$-dimensional faces, we get a linear transformation $\partial_n : C_n(K) \rightarrow C_{n-1}(K)$. Formally, for a simplex $\{p_0, \ldots, p_n\}$, let $\{p_0, \ldots, \hat{p}_i, \ldots, p_n\}$ denote the $(n-1)$-simplex obtained by omitting the vertex $p_i$.

**Definition 2.8.** The boundary operator $\partial_n : C_n(K) \rightarrow C_{n-1}(K)$ is the linear transformation defined on simplices by

$$\partial_n(\{p_0, \ldots, p_n\}) = \sum_{i=0}^n \{p_0, \ldots, \hat{p}_i, \ldots, p_n\}$$

The following lemma is what makes the theory work. Intuitively, it states that a boundary has no boundary.

**Lemma 2.9.** The composition $\partial_n \circ \partial_{n+1} = 0$ for all $n \geq 0$.

**Proof.** By linearity it suffices to shown that $\partial_n \circ \partial_{n+1}(\sigma) = 0$ for a $(n+1)$-simplex $\sigma = \{p_0, \ldots, p_{n+1}\}$.

$$\partial_n \circ \partial_{n+1}(\sigma) = \partial_n \left( \sum_{i=0}^{n+1} \{p_0, \ldots, \hat{p}_i, \ldots, p_{n+1}\} \right) = \sum_{i=0}^{n+1} \partial(\{p_0, \ldots, \hat{p}_i, \ldots, p_{n+1}\})$$

$$= \sum_{i=0}^{n+1} \sum_{j=0, j \neq i}^{n+1} \{p_0, \ldots, \hat{p}_j, \ldots, \hat{p}_i, \ldots, p_{n+1}\}.$$

For $i \neq j$ we see that the $(n-1)$-simplex $\{p_0, \ldots, \hat{p}_j, \ldots, \hat{p}_i, \ldots, p_{n+1}\}$ appears precisely two times in the sum. Since we are working over $\mathbb{Z}_2$, the result follows. \qed

We say that $c \in C_n(K)$ is an $n$-cycle if $\partial_n(c) = 0$ and we denote the associated vector space of $p$-cycles by $Z_n(K) = \ker \partial_n = \{c \in C_n(K) : \partial_n(c) = 0\}$. If $c$ is in the image of $\partial_{n+1}$, that is $c = \partial_{n+1}(d)$ for some $(n+1)$-cycle $d$, then $c$ is an $n$-boundary. We denote the associated vector space of $n$-boundaries by $B_n(K) = \text{Im} \partial_{n+1} = \{\partial_{n+1}(d) : d \in C_{n+1}(K)\}$.

**Example 2.10.** Consider the simplicial complex in Fig. 6. Observe that the cycle $c$ to the left can be written as a linear combination $c = c_1 + c_2 + c_3$. In fact, it is not hard to see that $\{c_1, c_2, c_3\}$ forms a basis for the cycles in $S_1(K)$: any “hole” in the simplicial complex can be expressed as a linear of those cycles. We are however interested in counting the number of “unbounded holes” and therefore we ignore any contribution from the cycles of $B_1(K) = \{c_3\}$. The number of “circular holes” is thus $\dim Z_1(K) - \dim B_1(K) = 3 - 1 = 2$.

With the previous example fresh in mind we define

**Definition 2.11.** The $n$-th Betti number of $K$ is $\beta_n(K) = \dim Z_n(K) - \dim B_n(K)$.

For simplicial complexes in $\mathbb{R}^3$ the interpretation of the Betti numbers is clear: $\beta_1$ is the number of loops, and $\beta_2$ is the number of cavities. In general, interpreting the betti numbers can be quite subtle. The following is however always true.
Definition 2.12. We say that a simplicial complex $K$ is **path-connected** if any two vertices in $K$ are connected by an edge path.

Remark 2.13. A simplicial complex $K$ is path-connected if and only if $|K|$ is path-connected in the topological sense. Proof: Assume that $|K|$ is path-connected. Fix a vertex $p$ in $K$ and let $K_0$ denote the simplicial complex (verify this) consisting of all simplices with the property that all of its vertices can be joined to $p$ by an edge-path. We let $K_1$ denote the subcomplex of $K$ consisting of all simplices that are not in $K_0$. Hence, $K = K_0 \cup K_1$ and $K_0 \cap K_1 = \emptyset$. Since $|K_0|$ and $|K_1|$ are closed subsets of $|K|$, it follows from connectivity that $K_1 = \emptyset$. The converse statement is immediate.

Lemma 2.14. $\beta_0(K)$ equals the number of path-connected components of $|K|$.

Proof. Observe that $Z_0(K) = S_0(K)$. Order the connected components of $|K|$ from 1 to $u$, and let $\{p^1_i, \ldots, p^u_i\}$ denote the set of vertices belonging to the $i$-th path-component. A basis for $Z_0(K)$ is given by $\bigcup_{i=1}^u \{p^1_i, p^2_i, \ldots, p^u_i, p^1_i\}$. Similarly, path-connectivity implies that there cannot be a 1-chain $c$ whose boundary is contained in $\{p^1_i, \ldots, p^u_i\}$. Hence, $\dim Z_0(K) - \dim B_1(K) = u$. $\square$

Theorem 2.15 (Euler-Poincare formula). For a simplicial complex $K$,

$$\chi(K) = \sum_{i=0}^{\infty} (-1)^i K_i = \sum_{i=0}^{\infty} (-1)^i \beta_i(K).$$

Proof. From linear algebra we remember that for a linear transformation $T: V \to W$ we have $\dim V = \dim \ker T + \dim \text{Im } T$. In our context this implies $K_i = \dim C_i(K) = \dim Z_i(K) + \dim B_{i-1}(K)$. Letting $B_{-1}(K) = 0$ we get the following:

$$\chi(K) = \sum_{i=0}^{\infty} (-1)^i K_i = \sum_{i=0}^{\infty} (-1)^i (\dim Z_i(K) + \dim B_{i-1}(K))$$

$$= \sum_{i=0}^{\infty} (-1)^i (\dim Z_i(K) - \dim B_i(K))$$

$$= \sum_{i=0}^{\infty} (-1)^i \beta_i(K).$$

$\square$
Computation  Computing $\beta_n(K)$ comes down to elementary linear algebra: represent $\partial_n$ in the canonical bases of $C_n(K)$ and $C_{n-1}(K)$ and reduce the resulting matrices to diagonal form using elementary row and column operations. From the diagonal form it is easy to read off $\dim B_{n-1}(K)$ as the number of non-zero diagonal terms, and $\dim Z_n(K)$ as the number of zero columns. The following example will make it clear.

**Example 2.16.** Let $K$ be the faces of a 2-simplex $\{1, 2, 3\}$. By identifying a simplex with its vertices and representing the matrices $\partial_2$ and $\partial_1$ in their standard bases we obtain

$$\partial_2 = \begin{pmatrix} 12 & 13 & 23 \\ 12 & 1 & 1 \\ 13 & -1 \\ 23 & 1 \end{pmatrix}$$

$$\partial_1 = \begin{pmatrix} 12 & 13 & 23 \\ 1 & 1 & 0 \\ 2 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix} \sim \text{row/column operations} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

We conclude that $\dim B_2(K) = 1$ and $\dim Z_1(K) = 1$, and therefore $\beta_1(K) = 1 - 1 = 0$.

**2.3 A Remark on the Choice of Coefficients**

The previous discussion has been for the field $k = \mathbb{Z}_2$. Homology is well-defined for coefficients in any other field, or more generally, in any *abelian group*. For simplicial complexes in $\mathbb{R}^3$ the Betti numbers are independent of the choice of coefficients. In higher dimensions however, the Betti numbers may depend on the underlying field. Instead of developing a full theory of simplicial homology in this more general setting, we shall limit ourselves to an example illustrating the subtleties arising.

**The boundary operator**  Let $k$ be any field and let $C_n(K; k)$ denote the free $k$-vector space on the $n$-simplices of $K$. In order for the boundary operator to satisfy $\partial_n \partial_{n+1} = 0$ the simplices need to be ordered. Intuitively, this means that we add a direction to every edge, and a left- or right-handed coordinate system to any 2-simplex, and so forth for higher dimensions. One way of obtaining such an orientation is to give a total order to the vertices of $K$:

$$p_0 < p_1 < \ldots < p_m.$$  

The boundary of the simplex $\{p_0, p_1\}$ is then $\partial_1(\{p_0, p_1\}) = p_1 - p_0$, and more generally, $\partial_n(\{p_0, \ldots, p_n\}) = \sum_{i=0}^{n} (-1)^i \{p_0, \ldots, \hat{p}_i, \ldots, p_n\}$. It is now a simple exercise to show that the boundary of the boundary vanishes. Hence, one gets cycle and boundary vector spaces satisfying $B_n(K; k) \subseteq Z_n(K; k)$.

**Example 2.17.** Let $K$ be the faces of a 2-simplex with vertex set $\{1, 2, 3\}$ and $k$ any field. By identifying a simplex with its vertices and representing the matrices $\partial_2$ and $\partial_1$ in their standard bases we obtain

$$\partial_2 = \begin{pmatrix} 12 & 13 & 23 \\ 12 & 1 & 1 \\ 13 & -1 \\ 23 & 1 \end{pmatrix}$$
Figure 7: projective plane

\[ \partial_1 = \begin{pmatrix} 12 & 13 & 23 \\ -1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} \sim \text{row/column operations} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \]

Hence the betti numbers are independent of the choice of coefficients.

**Example 2.18.** The projective plane \( \mathbb{R}P^2 \) is an unoriented closed surface which does not admit an embedding in \( \mathbb{R}^3 \). A triangulation of \( \mathbb{R}P^2 \) can be seen in Fig. 7. The diligent reader may write down the associated boundary matrices and verify that the Betti numbers differ between the choices \( k = \mathbb{Z}_3 \) and \( k = \mathbb{Z}_2 \):

\[
\begin{align*}
\beta_0(\mathbb{R}P^2; \mathbb{Z}_3) &= 1 & \beta_1(\mathbb{R}P^2; \mathbb{Z}_3) &= 0 & \beta_2(\mathbb{R}P^2; \mathbb{Z}_3) &= 0 \\
\beta_0(\mathbb{R}P^2; \mathbb{Z}_2) &= 1 & \beta_1(\mathbb{R}P^2; \mathbb{Z}_2) &= 1 & \beta_2(\mathbb{R}P^2; \mathbb{Z}_2) &= 1.
\end{align*}
\]

The Euler characteristic is obviously independent of the field.

### 2.4 Homology

The modern approach to defining the betti numbers is by means of vector spaces. Working with vector spaces, rather than mere integers, allow us to push homological information along continuous maps. This property, called functoriality, is instrumental in showing invariance under topological invariance and in the development of *persistent homology*.

**Definition 2.19.** The *n*-th simplicial homology vector space of a simplicial complex \( K \) is the quotient space \( H_n(K) = Z_n(K)/B_n(K) \).

It follows from the definition that \( \beta_n(K) = \dim H_n(K) \). We say that two cycles \( z_1 \) and \( z_2 \) are *homologous* if \( z_1 = z_2 + b \) where \( b \) is a boundary. We write \([z_1] = [z_2]\) for the associated equivalence class in \( H_n(K) \). Intuitively, two homologous cycles represent the same "hole" in the sense that their sum is either trivial or cycles around boundaries.
Lemma 2.20. Let $V \subseteq W$ be two finite-dimensional vector spaces (over some field). A basis $B'$ for $V$ extends to a basis $B$ for $W$. Furthermore, the elements of $B \setminus B'$ are representatives of a basis for $W/V$.

Proof. Left to the reader. \qed

The previous lemma provides us with an algorithm for computing a basis for $H_n(K)$: first find a basis $B'$ for $B_n(K)$ and then extend this basis to a basis $B$ for $Z_n(K)$. Then the elements of $B \setminus B'$ provide us with a basis for $H_n(K)$. Doing so by hand however can involve a significant amount of row reductions even for rather small simplicial complexes. Therefore we shall let the computer do the job!

Example 2.21. We return to the simplicial complex depicted in Fig. 6. The vector space $H_0(K)$ is generated by $[p_i]$ for any vertex $p_i$. Furthermore, $H_i(K) = 0$ for any $i > 1$. We have already seen that $B_2 = \{\sigma_3\}$. Hence $\{[c_1], [c_2]\}$ is a basis for $H_1(K) \cong \mathbb{Z}_2^2$. 

14
3 Week 3: Computation, Induced Maps and Invariance

This week will discuss the "standard algorithm" for computing simplicial homology, as well as induced maps and invariance of homology.

3.1 The Standard Algorithm

Denote the $i$-th column of a matrix $M$ by $M_i$. If $M_i$ is non-zero, then we define $\text{low}(M_i) = \max\{j : M_{ji} \neq 0\}$. For a zero-column $M_i$ we let $\text{low}(M_i) = 0$. We say that $M$ is reduced if $i \neq j$ implies $\text{low}(M_i) \neq \text{low}(M_j)$ or $0 = \text{low}(M_i) = \text{low}(M_j)$.

The following algorithm reduces the matrix $M$ to a reduced matrix $R$ with $R = MV$ where $V$ is a upper-triangular and of full rank.

**Data:** $m_1 \times m_2$ matrix $M$.

**Result:** Reduced $m_1 \times m_2$ matrix $R$, full-rank upper-triangular $V$, such that $R = MV$.

Let $R = M$ and $V = I_{m_2}$ (the $m_2 \times m_2$ identity matrix).

for $j \in \{1, 2, \ldots, m_2\}$ do

while there exists $i < j$ with $0 \neq \text{low}(R_i) = \text{low}(R_j)$ do

add column $i$ of $R$ to column $j$ of $R$;

add column $i$ of $V$ to column $j$ of $V$;

end

end

Algorithm 1: Reducing a matrix $M$ in $O((m_2)^3)$.

3.1.1 The Homology of a Triangle

We illustrate the algorithm on the simplicial complex given by the boundary of a triangle, $K = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}\}$. The first step is to stack all the boundary operators together in one big boundary matrix $D$ and define a corresponding identity matrix $V$:

$$R := D = \begin{pmatrix}
1 & 2 & 3 & 12 & 13 & 23 \\
1 & 0 & 0 & 0 & 1 & 1 & 0 \\
2 & 0 & 0 & 0 & 1 & 0 & 1 \\
3 & 0 & 0 & 0 & 0 & 1 & 1 \\
12 & 0 & 0 & 0 & 0 & 0 & 0 \\
13 & 0 & 0 & 0 & 0 & 0 & 0 \\
23 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}, \quad V = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.$$ 

Observe that $D$ is a linear transformation $D : C_0(K) \oplus C_1(K) \to C_0(K) \oplus C_1(K)$ and that we may identify every simplex with its corresponding basis vector in $C_0(K) \oplus C_1(K)$ in a natural way, i.e. the column vector $(100000)^T$ corresponds to the simplex $\{1\}$, $(000100)^T$ corresponds to $\{1, 2\}$ and $(000111)^T$ corresponds to the linear combination $\{1, 2\} + \{1, 3\} + \{2, 3\}$. Likewise, we see that $\ker D = Z_0(K) \oplus Z_1(K)$ and $\text{Im} D = B_0(K) \oplus B_1(K)$ as subspaces of $C_0(K) \oplus C_1(K)$.

We will now bring $D$ into reduced form by adding columns from left to right, and at the same time adding the corresponding columns in $V$. By iterating from the left to right, we check for every column $j$, if there exists a non-zero column $i < j$ with $\text{low}(R_i) = \text{low}(R_j)$. By adding column 5 to column 6, and then adding column 4 to column 6, we obtain the following two
matrices

\[
R := D = \begin{pmatrix}
1 & 2 & 3 & 12 & 13 & 13 + 12 + 23 \\
1 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & 1 & 0 \\
12 & 0 & 0 & 0 & 0 & 0 & 0 \\
13 & 0 & 0 & 0 & 0 & 0 & 0 \\
23 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}, \quad V = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.
\]

where \( R = DV \). The non-zero columns of a reduced matrix must necessarily be linearly independent, and therefore the non-zero columns corresponding to elements of \( C_1(K) \) will define a basis for \( B_0(K) \). Indeed, following the notation introduced above we see that \( \{(110000)^T, (101000)^T\} \) forms a basis for \( B_0(K) \). The zero-vectors necessarily form a basis for \( Z_0(K) \oplus Z_1(K) \). By construction, the \( i \)-th column of \( V \) corresponds precisely to the linear combination of simplices whose images under \( D \) adds up to the \( i \)-th column of \( R \). In other words, a basis for \( Z_0(K) \oplus Z_1(K) \) is given by the columns of \( V \) corresponding to 0-vectors in \( R \). Hence, \( \{(100000)^T, (010000)^T, (001000)^T\} \cup \{(000111)^T\} \) constitutes a basis for \( Z_0(K) \oplus Z_1(K) \). Since \( B_1(K) = 0 \) it follows that \( H_1(K) \cong \mathbb{Z}_2 \) with a basis \( \{(000111)^T\} \). Next step is to extend the basis for \( B_0(K) \) into a basis for \( Z_0(K) \), and then use Lemma 2.20 to obtain a basis for \( H_0(K) \).

Consider the two bases:

\[
\Sigma_{B_0} = \{(110000)^T, (101000)^T\},
\Sigma_{Z_0} = \{(100000)^T, (010000)^T, (001000)^T\}
\]

Since \( \text{low}((100000)^T) \not\in \{\text{low}((110000)^T), \text{low}((101000)^T)\} \), the set \( \Sigma_{B_0} \cup \{(100000)^T\} \) must necessarily be linearly independent. (This can be seen by constructing a matrix with the three vectors as column vectors and observing that the resulting matrix is reduced.) Furthermore, it contains the same number of basis elements as \( \Sigma_{Z_0} \) which spans \( Z_0(K) \). Hence, \( \Sigma_{Z_0} := \Sigma_{B_0} \cup \{(100000)^T\} \) forms a basis for \( Z_0(K) \) and \( H_0 \cong \mathbb{Z}_2 \) has a basis \( \{(100000)^T\} \).

### 3.1.2 The Algorithm

Let \( K \) be a simplicial complex of dimension \( d \), and let

\[
D : C_0(K) \oplus \ldots \oplus C_d(K) \to C_0(K) \oplus \ldots \oplus C_d(K)
\]

denote the boundary operator represented in the standard basis given by the simplices of all dimensions.

1. Use Algorithm [] to obtain a decomposition \( R = DV \).
2. The columns \( R_i \) with \( \text{low}(i) \neq 0 \) form a basis \( \Sigma_B = \Sigma_{B_0} \cup \ldots \cup \Sigma_{B_d} \) for \( B_0(K) \oplus \ldots \oplus B_d(K) \),
3. The columns \( V_i \) such that \( R_i = 0 \) form a basis \( \Sigma_Z = \Sigma_{Z_0} \cup \ldots \cup \Sigma_{Z_d} \) for \( Z_0(K) \oplus \ldots \oplus Z_d(K) \),
4. Let \( \Sigma_E = \Sigma_Z \setminus \{V_i \in \Sigma_Z : \exists R_j \text{ with } \text{low}(j) = \text{low}(i) = i\} \).
    (a) \( \Sigma_B \cup \Sigma_E \) is a basis for \( Z_0(K) \oplus \ldots \oplus Z_d(K) \),
    (b) \( \{[V_i] : V_i \in \Sigma_E\} \) is a basis for \( H_0(K) \oplus \ldots \oplus H_d(K) \).

Although the steps are justified in the example above, we briefly re-iterate the explanations.
1. 
2. The columns are linearly independent and all other columns are 0. Hence they must span the image of the linear transformation.
3. The zero-columns of $R$ correspond to the kernel of $D$. By construction, the basis elements for the kernel can be found in the corresponding columns of $V$.
4. (a) From the correctness of Algorithm 1, $\text{low}(R_i) \neq \text{low}(R_j)$ for any $R_i \neq R_j \in \Sigma_B$, and $\text{low}(V_i) \neq \text{low}(V_j)$ for any $V_i \neq V_j \in \Sigma_Z$. Since $\Sigma_B$ spans a subset of $\Sigma_Z$ in $C_0(K) \oplus \ldots \oplus C_d(K)$, there must exist exactly $|\Sigma_Z| - |\Sigma_B|$ vectors $V_i$ in $\Sigma_Z$ such that $\text{low}(V_i) \neq \text{low}(R_j)$ for all $R_j \in \Sigma_B$.
   (b) By Lemma 2.20

### 3.2 Simplicial Maps and their Induced Maps in Homology

Let $K$ and $L$ be abstract simplicial complexes with vertex sets $V_K$ and $V_L$ respectively. A **simplicial map** $f: K \rightarrow L$ is a function such that for every simplex $\{p_0, \ldots, p_n\}$ in $K$ its image $\{f(p_0), \ldots, f(p_n)\}$ is a simplex in $L$. Associated to a simplicial map we get an induced linear map in all dimensions:

$$ f_\#: C_n(K) \rightarrow C_n(L), \quad f_\#(\sigma) = \begin{cases} f(\sigma) & \text{if } \text{dim } f(\sigma) = \text{dim } \sigma \\ 0 & \text{otherwise} \end{cases}. $$

**Example 3.1.** Let $K$ denote the 2-simplex defined on the vertices $\{a, b, c\}$ and let $L = K - \{a, b, c\}$. Then $f: K \rightarrow L$ defines a simplicial map by $f(a) = a$, $f(b) = f(c) = b$. It follows that $f_\#: C_2(K) \rightarrow C_2(L)$ is the zero-map, and that $f_\#: C_1(K) \rightarrow C_1(L)$ has matrix representation

$$
\begin{pmatrix}
ab & ac & bc \\
ab & 1 & 1 & 0 \\
ac & 0 & 0 & 0 \\
bc & 0 & 0 & 0
\end{pmatrix}.
$$

**Lemma 3.2.** We have the following commutative diagram

$$
\begin{array}{ccc}
C_n(K) & \xrightarrow{\partial_n} & C_{n-1}(K) \\
| f_\# | & & | f_\# |
\end{array}
\begin{array}{ccc}
C_n(L) & \xrightarrow{\partial_n} & C_{n-1}(L)
\end{array}
$$

I.e., $f_\# \circ \partial_n = \partial_n \circ f_\#$.

**Proof.** It suffices to prove the result for a single simplex $\sigma = \{p_0, \ldots, p_n\}$.

**Case 1:** $\text{dim } f(\sigma) = \text{dim } \sigma$. In this case we have that

$$ f_\#(\partial_n(\sigma)) = \sum_{i=0}^{n} f_\#(\{p_0, \ldots, \hat{p}_i, \ldots, p_n\}) = \sum_{i=0}^{n} \{f(p_0), \ldots, \hat{f(p}_i), \ldots, f(p_n)\} = \partial_n \circ f_\#(\sigma). $$

**Case 2:** $\text{dim } f(\sigma) \leq \text{dim } \sigma - 2$. It is immediate that $\partial_n \circ f_\#(\sigma) = 0$. Furthermore, $\text{dim}(\{f(p_0), \ldots, \hat{f(p)_i}, \ldots, f(p_n)\}) \leq \text{dim } \sigma - 2$, and therefore

$$\sum_{i=0}^{n} f_\#(\{p_0, \ldots, \hat{p}_i, \ldots, p_n\}) = \sum_{i=0}^{n} 0 = 0.$$
Case 3: \( \dim f(\sigma) = \dim \sigma - 1 \). It is immediate that \( \partial_n \circ f_\#(\sigma) = 0 \). Assume without loss of generality that \( f(p_0) = f(p_1) \). Then \( \dim(\{f(p_0), \ldots, f(p_i), \ldots, f(p_n)\}) \leq \dim \sigma - 2 \) for all \( 2 \leq i \leq n \). Thus,
\[
\sum_{i=0}^{n} f_\#(\{p_0, \ldots, \hat{p}_i, \ldots, p_n\}) = f_\#(\{p_1, \ldots, p_n\}) + f_\#(\{p_0, \ldots, p_n\}) = 0.
\]

**Corollary 3.3.** A simplicial map \( f: K \to L \) induces a well-defined map \( f_*: H_n(K) \to H_n(L) \) defined by \( f_*(\langle c \rangle) = [f_\#(c)] \).

**Proof.** Since \( H_n(K) = Z_n(K)/B_n(K) \) and \( H_n(L) = Z_n(L)/B_n(L) \), it suffices to show that \( f_\#(Z_n(K)) \subseteq Z_n(L) \) and \( f_\#(B_n(K)) \subseteq B_n(L) \).

Assume that \( \partial_n(c) = 0 \) for \( c \in C_n(K) \). Then by Lemma 3.2 we have \( \partial_n(f_\#(c)) = f_\#(\partial_n(c)) = 0 \). We conclude that \( f_\#(c) \in Z_n(L) \). Now assume that \( c = \partial_{n+1}(d) \) for some \( d \in C_{n+1}(K) \). Then \( f_\#(c) = f_\#(\partial_{n+1}(d)) = \partial_{n+1}(f_\#(d)) \). Hence \( f_\#(c) \in B_n(L) \). \( \square \)

### 3.3 Simplicial Approximations and Invariance of Homology

In order to put the homology vector spaces on a firm footing we need to prove that if \( |K| \) is homeomorphic to \( |L| \), then \( H_n(K) \cong H_n(L) \). A complete proof of this fact is no easy task and we will not discuss the proof in detail. The goal of this part is merely to outline how a continuous map \( h: |K| \to |L| \) induces a map \( h_*: H_n(K) \to H_n(L) \) in homology, and that this construction is functorial. An easy consequence will be that \( f_* \) is an isomorphism whenever \( f \) is a homeomorphism. The full story can be found in [10].

#### 3.3.1 Simplicial Maps Revisited

For geometric simplicial complexes \( K \) and \( L \) we define a simplicial map \( f: K \to L \) to be a function \( f: K^{(0)} \to L^{(0)} \) such that if \( \sigma \) is spanned by \( p_0, \ldots, p_n \), then the simplex spanned by \( f(p_0), \ldots, f(p_n) \) is contained in \( L \).

**Remark 3.4.** Recall that any geometric simplicial complex \( K \) can be naturally identified with an abstract simplicial complex through its vertex scheme, i.e. by identifying every geometric simplex with the vertices that span it. A simplicial map of geometric simplicial complexes naturally induces a simplicial map between the associated vertex schemes. Hence, a simplicial map \( f: K \to L \) between geometric simplicial complexes induces a map \( f_*: H_n(K) \to H_n(L) \) in homology.

The simplicial map \( f: K \to L \) induces a continuous (affine) map \( |f|: |K| \to |L| \) given by
\[
|f| \left( \sum_{i=0}^{n} t_i p_i \right) = \sum_{i=0}^{n} t_i f(p_i), \quad \text{where} \quad \sum_{i=0}^{n} t_i = 1, \quad \text{and} \quad 0 \leq t_i \leq 1.
\]

#### 3.3.2 Simplicial Approximations

We shall now see how a well-defined morphism \( h_*: H_n(K) \to H_n(L) \) can be associated to any continuous map \( h: |K| \to |L| \).

**Definition 3.5.** Let \( K \) and \( L \) be geometric simplicial complexes and \( h: |K| \to |L| \). We say that \( f: K \to L \) is a simplicial approximation to \( h \), if for each \( x \in |K| \), \( |f|(x) \) is contained in the minimal simplex which contains \( h(x) \).
The following theorem makes it clear why simplicial approximations are relevant.

**Theorem 3.6.** If $f_1, f_2: K \to L$ are simplicial approximations to $h: |K| \to |L|$, then $f_{1,*} = f_{2,*}: H_n(K) \to H_n(L)$.

In particular, for any continuous map $h$ that admits a simplicial approximation $f$, we have a well-defined map $f_*: H_n(K) \to H_n(L)$ in homology. However, not all maps admit such approximations.

**Example 3.7.** Let $K$ be the triangulation of the triangle shown in Fig. 8 and let $L$ be the triangulation of a filled rectangle with a hole in the same figure. The curve in $|L|$ shows the image of $|K|$ under a continuous mapping $h$. We see that $h$ admits a simplicial approximation $f: K \to L$ which is obtained by pushing the images of the vertices of $K$ onto the corresponding vertices of $L$ as indicated by arrows. Note that if $|K|$ had a less refined triangulation given by $\{\{a\}, \{b\}, \{c\}, \{a,b\}, \{a,c\}, \{b,c\}\}$ then $h$ would not admit a simplicial approximation.

The previous example shows that for a given $K$ and $L$, and a function $h: |K| \to |L|$, there need not exist a simplicial approximation $f: K \to L$ to $h$. However, it seems plausible that one can always "subdivide" the simplices of $K$ to be so "small" that we can push the image of its vertices onto the vertices of $L$ in a way that gives a well-defined simplicial map.

**Definition 3.8.** A simplicial complex $K_1$ is a **subdivision** of a simplicial complex $K$ if $|K_1| = |K|$ and each simplex of $K_1$ is contained in a simplex of $K$.

One way to subdivide a simplicial complex $K$ is through **barycentric subdivision**. Let $\sigma \in K$ be a simplex in $\mathbb{R}^k$ spanned by vertices $\{p_0, \ldots, p_n\}$. The barycenter of $\sigma$ is

$$\hat{\sigma} = \frac{1}{n+1} (p_0 + \ldots + p_n).$$

The first **barycentric subdivision** of $K$ is the simplicial complex $\text{Sd}(K)$ consisting of all simplices spanned by vertices $\hat{\sigma}_0, \ldots, \hat{\sigma}_r$ where $\sigma_0 \subset \sigma_1 \subset \ldots \subset \sigma_r$ is a sequence of different simplices in $K$. In particular, the vertex set of $\text{Sd}(K)$ is $\{\hat{\sigma} : \sigma \in K\}$. 

---

**Figure 8: Simplicial approximation**
Remark 3.9. $\text{Sd}(K)$ is a simplicial complex and $|\text{Sd}(K)| = |K|$.

For $j > 1$ we define the $j$-th barycentric subdivision $\text{Sd}^j(K)$ to be the barycentric subdivision of $\text{Sd}^{j-1}(K)$.

Example 3.10. Fig. 9 shows a simplicial complex $K$ together with its first two barycentric subdivisions $\text{Sd}(K)$ and $\text{Sd}^2(K)$. Observe that any function $f: \text{Sd}(K)^{(0)} \rightarrow K^{(0)}$ which maps a vertex $\hat{\sigma}$ to any vertex of $\sigma$ defines a simplicial map.

We state the following important theorem due to Alexander without a proof.

Theorem 3.11 (Simplicial Approximation Theorem). Let $K$ and $L$ be simplicial complexes and $h: |K| \rightarrow |L|$ a continuous map. Then $h$ admits a simplicial approximation $f: \text{Sd}^j(K) \rightarrow L$ for some sufficiently large $j$.

In conjunction with Theorem 3.6 this shows that $h$ induces a well-defined map $f_*: H_n(\text{Sd}^j(K)) \rightarrow H_n(L)$ for some sufficiently large $j$.

Lemma 3.12. Let $f: \text{Sd}(K)^{(0)} \rightarrow K^{(0)}$ denote any map which sends a vertex $\hat{\sigma}$ in $\text{Sd}(K)$ to any vertex of $\sigma$. Then $f$ is a simplicial map and an approximation to the identity $\text{id}_{|K|}: |\text{Sd}(K)| \rightarrow |K|$.

Hence, Theorem 3.6 implies that any such map induces a canonical map $H_n(\text{Sd}(K)) \rightarrow H_n(K)$. Crucially, it turns out that this morphism is in fact an isomorphism. Let $\eta_j: H_n(\text{Sd}^j(K)) \rightarrow H_n(K)$ denote the canonical isomorphism given by the composition

$$H_n(\text{Sd}^j(K)) \xrightarrow{\cong} H_n(\text{Sd}^{j-1}) \xrightarrow{\cong} \cdots \xrightarrow{\cong} H_n(\text{Sd}(K)) \xrightarrow{\cong} H_n(K).$$

Theorem 3.13. Any continuous map $h: |K| \rightarrow |L|$ between polyhedra of simplicial complexes induces a well-defined linear map $h_* := f_* \circ \eta_j^{-1}: H_p(K) \rightarrow H_p(L)$ in homology, where $f: \text{Sd}^j(K) \rightarrow L$ is a simplicial approximation to $h$ and $\eta_j: H_n(\text{Sd}^j(K)) \rightarrow H_n(K)$ is the canonical isomorphism. Furthermore, $h_*$ is functorial, i.e. $(\text{id}_{|K|})_* = \text{id}_{H_n(K)}$ and $(g \circ h)_* = g_* \circ h_*$ for any pair of continuous maps $h: |K| \rightarrow |L|$ and $g: |L| \rightarrow |M|$.

If $h: |K| \rightarrow |L|$ is a homeomorphism, then

$$\text{id}_{H_n(K)} = (\text{id}_{|K|})_* = (h^{-1} \circ h)_* = (h^{-1})_* \circ h_*: H_n(K) \rightarrow H_n(K),$$

$$\text{id}_{H_n(L)} = (\text{id}_{|L|})_* = (h \circ h^{-1})_* = h_* \circ (h^{-1})_*: H_n(L) \rightarrow H_n(L).$$

Hence, $h_*$ is an isomorphism.
4 Week 4: Homotopy Equivalence, Relative Homology and Sensor Networks

4.1 Homotopy Equivalence

In the last lecture we outlined how a continuous map \( h: |K| \to |L| \) between polyhedra induces a linear map in homology, and that this map is an isomorphism if \( h \) is a homeomorphism. We shall now discuss a weaker condition on \( h \) for which the induced map is an isomorphism. Every function is assumed to be continuous.

**Definition 4.1.** A homotopy between two maps \( f, g: X \to Y \) is a map \( F: X \times [0, 1] \to Y \), such that \( F(x, 0) = f(x) \) and \( F(x, 1) = g(x) \) for all \( x \in X \). (\( X \times [0, 1] \) is equipped with the product topology.)

If there is a homotopy between \( f \) and \( g \) then we say that \( f \) and \( g \) are homotopic and denote this by \( f \simeq g \). It is a standard fact that homotopy defines an equivalence relation.

**Example 4.2.** Fig. 10 shows an example of a homotopy between two functions \( f, g: [0, 1] \to \mathbb{R}^2 \).

**Example 4.3.** The two functions \( f, g: [0, 1] \to \{0, 1\} \) defined by \( f(x) = 0 \) and \( g(x) = 1 \) for all \( x \) are not homotopic.

We will however need a stricter version of homotopy. Let \( A \subseteq X \) and \( B \subseteq Y \). A map of pairs \( f: (X, A) \to (Y, B) \) is a map \( f: X \to Y \) such that \( f(A) \subseteq B \). A homotopy of pairs between \( f, g: (X, A) \to (Y, B) \) is a homotopy \( F: X \times [0, 1] \to Y \) between \( f \) and \( g \) such that \( F_t(A) \subseteq B \) for all \( t \in [0, 1] \).

**Example 4.4.** Let \( X = [0, 1] \), \( A = \{0, 1\} \), \( Y = \mathbb{R}^2 - \{0\} \) and \( B = \{p\} \) where \( 0 \neq p \in \mathbb{R}^2 \). The equivalence classes of homotopic maps constitute the fundamental group of \( \mathbb{R}^2 - \{0\} \) (w.r.t point \( p \)); see Fig. 11.

**Definition 4.5.** The spaces \((X, A)\) and \((Y, B)\) are homotopy equivalent if there exists maps \( f: (X, A) \to (Y, B) \) and \( g: (Y, B) \to (X, A) \) such that \( g \circ f \simeq \text{id}_{(X, A)} \) and \( f \circ g \simeq \text{id}_{(Y, B)} \).

We write \((X, A) \simeq (Y, B)\) if \((X, A)\) and \((Y, B)\) are homotopy equivalent. If \( i: (Y, B) \to (X, A) \) denotes an inclusion of a subspace, and \( r: (X, A) \to (Y, B) \) is such that \( r \circ i = \text{id}_{(Y, B)} \) and \( i \circ r \simeq \text{id}_{(X, A)} \), then \((Y, B)\) is a deformation retract of \((X, A)\). If \((X, A) \simeq \{pt\}\) then \((X, A)\) is contractible; see Fig. 12.

Our interest in homotopic maps stems from the following theorem.

**Theorem 4.6.** If \( f \simeq g: |K| \to |L| \) then \( f_* = g_*: H_n(K) \to H_n(L) \).

**Corollary 4.7.** If \( f: |K| \to |L| \) is a homotopy equivalence of polyhedra, then is \( f_*: H_n(K) \to H_n(L) \) an isomorphism.
Figure 11: The functions $f$ and $g$ are homotopic, but neither of them is homotopic to $h$.

Figure 12: The disk deformation retracts down onto its origin and is thus contractible. The disk with a hole in the middle deformation retracts onto the circle (and is not contractible).
In conclusion: from time to time we need to consider the homology of a topological space which is homotopy equivalent to a polyhedra. While this is (in a technical sense) possible when restricted to triangulated spaces, the formal treatment is somewhat technical and we will not consider such details.

There is a more general homology theory called singular homology which is defined for any topological space and which coincides with simplicial homology when restricted to polyhedra. So in particular, whenever \( X \simeq |K| \), we can compute its singular homology \( H^S_n(X) \) by means of the simplicial homology \( H_n(K) \). Furthermore, if \( f: X \to Y \) with \( |K| \simeq X \) and \( Y \simeq |L| \) then there is the following commutative diagram where \( \hat{f} \) is the map given by the composition \( |K| \to X \xrightarrow{f} Y \to |L| \),

\[
\begin{array}{ccc}
H^S_n(X) & \xrightarrow{f_*} & H^S_n(Y) \\
\downarrow\cong & & \downarrow\cong \\
H_n(K) & \xrightarrow{\hat{f}_*} & H_n(L).
\end{array}
\]

In conclusion: from time to time we need to consider the homology of a topological space \( X \) which is not a polyhedron and we will simply denote its homology by \( H_n(X) \) where this is to be understood as the simplicial homology of any polyhedron homotopy equivalent to \( X \). We will not consider spaces which are not homotopy equivalent to a simplicial complex.

### 4.2 The Nerve Theorem

A particular instance of a homotopy equivalence is the nerve theorem.

Let \( F \) be a finite collection of closed, convex sets in Euclidean space. Define the nerve of \( F \) to be the abstract simplicial complex

\[
\mathcal{N}(F) := \{G \subseteq F: \bigcap_{F_i \in G} F_i \neq \emptyset\}.
\]

**Example 4.9.** Consider the set \( F = \{F_1, F_2, F_3\} \) in Fig. 13. Then

\[
\mathcal{N}(F) = \{\{F_1\}, \{F_2\}, \{F_3\}, \{F_1, F_2\}, \{F_2, F_3\}, \{F_1, F_3\}\}
\]

which is (isomorphic to) the abstract simplicial complex associated to the boundary of a 2-simplex. Clearly, \( F_1 \cup F_2 \cup F_3 \simeq |\mathcal{N}(F)| \).

If \( F = \{F_1, \ldots, F_n\} \) and \( F' = \{F'_1, \ldots, F'_m\} \) are such that \( F_i \subseteq F'_i \) for all \( F_i \in F \), then there is an induced simplicial map \( \sigma: \mathcal{N}(F) \to \mathcal{N}(F') \) defined by sending \( F_i \) to \( F'_i \). In particular, we have an induced linear transformation \( \sigma_*: H_n(\mathcal{N}(F)) \to H_n(\mathcal{N}(F')) \).

**Theorem 4.10.** 1. Let \( F \) be as above, then \( \bigcup_{F_i \in F} F_i \simeq |\mathcal{N}(F)| \).
2. Let $F$ and $F'$ be as above and let $j: \bigcup_{F_i \in F} F_i \hookrightarrow \bigcup_{F'_i \in F'} F'_i$ denote the inclusion. Then the following diagram commutes:

$$
\begin{array}{c}
H_n(\bigcup_{F_i \in F} F_i) \xrightarrow{j_*} H_n(\bigcup_{F'_i \in F'} F'_i) \\
\cong \downarrow \quad \cong \\
H_n(N(F)) \xrightarrow{\sigma_*} H_n(N(F'))
\end{array}
$$

4.2.1 The Cech Complex

Let $P \subset \mathbb{R}^n$ be a finite set of points. By definition of the Cech complex we have that $\text{Cech}_r(P) = \mathcal{N}(F)$ where $F = \{B_r(p): p \in P\}$, and it follows from the Nerve theorem that the geometric realization of the Cech complex is homotopy equivalent to the union of closed $r$-balls. This shows that the Betti numbers associated to the Cech complex can be interpreted as "holes" of various dimension formed by the associated union of balls in Euclidean space; see Fig. 14.

4.3 Relative Homology

Let $K_0 \subseteq K$ be a subcomplex. The vector space of relative $n$-chains is the quotient vector space $C(K,K_0) = C(K)/C(K_0)$. Since $\partial_n(C_n(K_0)) \subseteq C_{n-1}(K_0)$ by construction, it follows that $\partial$ descends to a boundary map

$$
\partial_n: C_n(K,K_0) \to C_{n-1}(K,K_0)
$$

such that $\partial_n \circ \partial_{n+1} = 0$. Hence, we get associated relative homology groups given by

$$
H_n(K,K_0) = \frac{Z_n(K,K_0)}{B_n(K,K_0)} := \frac{\ker \partial_n: C_n(K,K_0) \to C_{n-1}(K,K_0)}{\text{Im} \partial_{n+1}: C_{n+1}(K,K_0) \to C_n(K,K_0)}.
$$
Example 4.11. Let $K$ denote the simplicial complex given by a 2-simplex and all its faces, and let $K_0$ denote its 1-skeleton. It follows that:

\[
\begin{align*}
\dim C_2(K) &= 1 \\
\dim C_1(K) &= 3 \\
\dim C_0(K) &= 3
\end{align*}
\]

We conclude that $H_2(K, K_0) \cong \mathbb{Z}_2$, $H_1(K, K_0) = H_0(K, K_0) = 0$.

The previous example generalizes to any $n$-simplex: let $K$ be denote a $n$-simplex with its faces, and $K_0$ its $(n-1)$-skeleton. Then $H_n(K, K_0) \cong \mathbb{Z}_2$ and all other homology vector spaces are trivial. The following theorem explains why.

**Theorem 4.12.** Let $K_0 \subseteq K$. Then, $H_n(K, K_0) \cong H_n([K]/[K_0])$ for all $n \geq 1$.

Here $[K]/[K_0]$ need not be a simplicial complex; see Remark 4.8.

Example 4.13. Returning to the previous example we see that $[K]/[K_0] \approx S^2 \simeq Bd(\Delta^3)$ (the boundary of the standard 3-simplex). Thus, $H_2(K, K_0) \cong H_2([K]/[K_0]) \cong H_2(Bd(\Delta^3)) \cong \mathbb{Z}_2$.

Next we observe that if $f : (K, K_0) \to (L, L_0)$ is a (geometric) simplicial map of pairs, then $f_\#(C_n(K_0)) \subseteq C_n(L_0)$. Therefore the induced map $f_\# : C_n(K, K_0) \to C_n(L, L_0)$ defined by $f_\#([c]) = [f(c)]$ is well-defined. Furthermore, it follows from Lemma 3.2 that

\[
\partial_n \circ f_\#([c]) = [\partial_n \circ f_\#(c)] = [f_\# \circ \partial_n(c)] = f_\# \circ \partial_n([c]).
\]

Just as in the proof of Corollary 3.3 this gives an induced map in relative homology

\[
f_* : H_n(K, K_0) \to H_n(L, L_0).
\]

**Theorem 4.14.** If $f, g : ([K], [K_0]) \to ([L], [L_0])$ are homotopic, then $f_* = g_* : H_n(K, K_0) \to H_n(L, L_0)$.

**Corollary 4.15.** If $h : ([K], [K_0]) \to ([L], [L_0])$ is a homotopy equivalence, then $h_*$ an isomorphism.

**Proof.** See the proof of Corollary 4.7.

Example 4.16. Let $D^2$ denote the unit disk. Then $(D^2 - \{0\}, S^1) \simeq (S^1, S^1)$ and therefore $H_2(D^2 - \{0\}, S^1) \cong 0$. On the other hand, $(D^2, S^1) \simeq (\Delta^2, Bd(\Delta^2))$ and therefore

\[
H_2(D, S^1) \cong H_2(\Delta^2/Bd(\Delta^2)) \cong H_2(S^2) \cong H_2(Bd(\Delta^3)) \cong \mathbb{Z}_2.
\]

The following linear transformation will be important in the next section.

**Theorem 4.17.** Let $K_0 \subseteq K$. The linear map $\delta_* : H_n(K, K_0) \to H_{n-1}(K_0)$ given by $\delta_*([c]) = [\partial_n(c)]$ is well-defined, and if $f : (K, K_0) \to (L, L_0)$ is a simplicial map of pairs, then the following diagram commutes:

\[
\begin{array}{ccc}
H_n(K, K_0) & \xrightarrow{\delta_*} & H_{n-1}(K_0) \\
\downarrow{f_*} & & \downarrow{f_*} \\
H_n(L, L_0) & \xrightarrow{\delta_*} & H_{n-1}(L_0).
\end{array}
\]

**Proof.** Left as an exercise.
4.4 Sensor Networks

We now show how relative homology can be used to determine if a region is covered by a finite set of simple sensors. We will prove \cite{6} Theorem 1. For a beautiful overview of homological sensor networks see \cite{7}.

4.4.1 The covering problem

The problem is as follows: assume that you are given a collection of nodes $X$ (i.e. a finite point set) contained in a domain $D \subset \mathbb{R}^2$, where each node can sense, broadcast, and cover a region of a given radius. Does the collection of coverage disks of $X$ cover the domain? In the illustration above we see that the answer is "yes" in the left case and "no" in the right case. Assuming the location of every sensor is known, then this is a trivial task. But we will work under the assumption that we do not known the specific location of the sensors - merely which sensors are close to each other. Again, see the above references for the motivation for these sensor networks.

We will work under four assumptions:

1. Nodes $X$ broadcast their unique ID numbers. Each node can detect the identity of any node within broadcast radius $r_b$.
2. Nodes have radially symmetric covering domains of cover radius $r_c \geq r_b/\sqrt{3}$.
3. Nodes $X$ lie in a compact connected domain $D \subset \mathbb{R}^2$ whose boundary $\partial D$ is connected and piecewise-linear with vertices marked fence nodes $X_f$.
4. Each fence node $x \in X_f$ knows the identities of its neighbors on $\partial D$ and these neighbors both lie within distance $r_b$ of $x$.

For simplicity we employ the notation $R := VR_{r_b/2}(X)$ and we let $F$ denote the 1-dimensional cycle associated to the boundary nodes, i.e. $\partial D$ is a geometric realization of $F$. Lastly, recall from Theorem \ref{thm:boundary} that there is a well-defined morphism $\delta_* : H_2(R, F) \to H_1(F)$.

**Theorem 4.18.** If there exists $[\alpha] \in H_2(R, F)$ such that $0 \neq \delta_*([\alpha]) = [\partial_2(\alpha)] \in H_1(F)$, then the union $U = \bigcup_{x \in X} B_{r_c}(x)$ covers $D$.

**Proof.** Let $z : |R| \to \mathbb{R}^2$ be the affine map which sends a vertex $v_i$ of $|R|$ to its corresponding node in the plane, and which extends linearly to all higher-dimensional simplices: $z(\sum_{i=0}^n t_i v_i) = \sum_{i=0}^n t_i z(v_i)$. First we observe that $z(|R|) \subseteq U$. Indeed, let $Q$ be a set of nodes which form a simplex $\tau$ in $R$, i.e. $\text{Diam } Q \leq r_b$. Then, by Theorem \ref{thm:boundary} $z(\tau)$ is contained in a closed ball with radius $r_b/\sqrt{3}$, or equivalently, $z(\tau)$ is contained in

$$\bigcup_{x \in Q} B_{r_b/\sqrt{3}}(x) \subseteq \bigcup_{x \in Q} B_{r_c}(x) \subseteq U.$$
Since this holds for every simplex $\tau$, it follows that $z(|R|) \subseteq U$.

From Theorem 4.17 we have the following commutative diagram:

$$
\begin{array}{ccc}
H_2(R, F) & \xrightarrow{\delta_*} & H_1(F) \\
\downarrow z_* & & \downarrow z_* \\
H_2(\mathbb{R}^2, \partial D) & \xrightarrow{\delta_*} & H_1(\delta D).
\end{array}
$$

Since the rightmost $z_*$ is an isomorphism, it follows that $z_* \circ \delta_*(|\alpha|) \neq 0$ and therefore it must be true that $z_*(|\alpha|) \neq 0$. Assume for the sake of contradiction that there is a $p \in D$ which is not contained in $U$. Then the map $z$ factors as

$$
\begin{array}{ccc}
(|R|, |F|) & \xrightarrow{z} & (\mathbb{R}^2, \partial D) \\
\downarrow z & & \uparrow \\
(U, \partial D) & \longrightarrow & (\mathbb{R}^2 \setminus \{p\}, \partial D)
\end{array}
$$

But then we get the following diagram in homology where the top arrow is non-zero,

$$
\begin{array}{ccc}
H_2(|R|, |F|) & \xrightarrow{z_*} & H_2(\mathbb{R}^2, \partial D) \\
\downarrow z_* & & \uparrow \\
H_2(U, \partial D) & \longrightarrow & H_2(\mathbb{R}^2 \setminus \{p\}, \partial D)
\end{array}
$$

This leads to a contradiction because $(\mathbb{R}^2 \setminus \{p\}, \partial D) \simeq (\partial D, \partial D)$ and thus

$$
H_2(D \setminus \{p\}, \partial D) \cong H_2(\partial D, \partial D) = 0.
$$

\[\square\]

**Corollary 4.19.** Let $\alpha$ be as in Theorem 4.18. Then the restriction of $U$ to those nodes which make up $\alpha$ covers $D$. 

27
5 Week 5: Topology & Homology Inference and Persistent Betti Numbers

5.1 Topology & Homology Inference

Given an (unknown) compact subset $K \subseteq \mathbb{R}^d$ and a set of points $P$ which lie ’close’ to $K$, is it possible to infer the topology of $K$ from $P$ assuming the sample is dense enough and with little noise? We shall state one theorem which answers that question to the affirmative under quite strong assumptions on $K$. These assumptions can be significantly weakened; See Chapter 4 and 5 of [11].

For $K$ and $P$ as above and $x \in \mathbb{R}^d$, define $d_K(x) = \min_{y \in K} ||x - y||$, $d_P(x) = \min_{y \in P} ||x - y||$ where the norm is the standard 2-norm. We get the following associated sublevel-set filtrations:

$K_r := d_K^{-1}(-\infty, r]$ and $P_r := d_P^{-1}(-\infty, r] = \cup_{p \in P} B_r(p)$.

Definition 5.1. The Hausdorff distance between two non-empty subsets $X, Y \subseteq \mathbb{R}^d$ is defined by $d_H(X, Y) := \max\{\sup_{x \in X} \inf_{y \in Y} ||x - y||, \sup_{y \in Y} \inf_{x \in X} ||y - x||\}$.

For a given $x \in \mathbb{R}^d$ we define $\Pi_K(x)$ to be the set of all points $y \in Y$ such that $d_K(x) = ||x - y||$. Note that $|\Pi_K(x)| \geq 1$ for all $x$. The medial axis of $K$ is the set $M(K) := \{x : |\Pi_K(x)| \geq 2\}$, see Fig. 15 for an illustration.

Definition 5.2. The reach of $K$ is given by

$Rch(K) := \min_{y \in K} d(y, \overline{M(K)})$

where $\overline{M(K)}$ denotes the topological closure of $M(K)$.

Remark 5.3. If $K$ has positive reach, then one can show (not surprisingly) that the map $K_r \to K$ defined by mapping $x$ to $\Pi_K(x)$ is a homotopy equivalence for all $r < Rch(K)$.

Theorem 5.4 (Niyogi, Smale, Weinberger (2006)). Let $K$ be a compact submanifold of $\mathbb{R}^d$ with positive reach. Let $P$ be a point set such that $d_H(K, P) = \epsilon < \sqrt{\frac{3}{50}} Rch(K)$. Then for any $r \in (2\epsilon, \sqrt{\frac{2}{5}} Rch(K))$, the offset $P_r$ deformation retracts onto $K$. 

28
This shows that the topology of $K$ can be recovered under fairly strong assumptions. By replacing the reach with the weak feature size these results can be generalized substantially, but that is beyond the scope of this course. In particular, for a "reasonable" space $K$ together with an adequate sampling, one can recover the homology $H_n(K)$ from $H_n(P_r)$ for some $r \geq 0$.

We now proceed to give an elementary proof of the fact that the homology of a (not necessarily compact) space can be recovered from the persistent homology of a sufficiently good sampling.

**Theorem 5.5.** Let $K \subseteq \mathbb{R}^d$, $P \subseteq \mathbb{R}^d$ a finite point set and $\epsilon, \delta > 0$ such that

1. $K \subseteq P_\delta$ (sampling density),
2. $P \subseteq K_\epsilon$ (sampling error: points need not lie on $K$),
3. The inclusion $K \hookrightarrow K_{\delta+\epsilon}$ induces an isomorphism $H_n(K) \to H_n(K_{\delta+\epsilon})$,
4. The inclusion $K_{\delta+\epsilon} \hookrightarrow K_{2(\epsilon+\delta)}$ induces a monomorphism $H_n(K_{\delta+\epsilon}) \to H_n(K_{2(\epsilon+\delta)})$.

Then $H_n(K) \cong \text{Im}(H_n(P_\delta) \to H_n(P_{2\delta+\epsilon}))$.

**Proof.** Assumption 1. and 2. yield the following commutative diagram of inclusions

$$
\begin{array}{ccc}
K & \to & K_{\delta+\epsilon} \\
\downarrow & & \downarrow \\
K_{2(\epsilon+\delta)} & \to & K_{2(\epsilon+\delta)} \\
\end{array}
$$

$$
\begin{array}{ccc}
P_\delta & \to & P_{2\delta+\epsilon} \\
\end{array}
$$

which induces the following commutative diagram in homology (see Remark 4.8 and Theorem 3.6)

$$
\begin{array}{ccc}
H_n(K) & \overset{\cong}{\to} & H_n(K_{\delta+\epsilon}) \\
\downarrow & & \downarrow \text{mono} \\
H_n(P_\delta) & \to & H_n(P_{2\delta+\epsilon}) \\
\end{array}
$$

Since the top-left morphism is an isomorphism, we must have that $a$ is surjective. Similarly, $b$ must be a monomorphism. We conclude that

$$
H_n(K) \cong H_n(K_{\delta+\epsilon}) \cong \text{Im } b = \text{Im } b \circ a = \text{Im}(H_n(P_\delta) \to H_n(P_{2\delta+\epsilon})).
$$

\square

### 5.2 Persistent Betti Numbers

For a point set $P \subseteq \mathbb{R}^d$, the Betti numbers $\beta_i(P_r)$ are very unstable in the scale parameter $r$. That is, a small perturbation of the point set may result in completely different Betti numbers. Furthermore, although there does exist a "perfect" scale for a good enough sampling of a sufficiently nice space, it is not at all clear what this scale actually is. Therefore we shall study a tool from applied algebraic topology which is both stable and scale-free.

Let us first consider $\beta_0$ - the connected components. In this setting things are particularly simple and we can use a tool from statistics called a **dendrogram**.

**Example 5.6.** Let $P = \{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9\}$ be the points of Fig. 16. When considered at a sufficiently small scale, $P_r$ consists of 9 connected components (clusters). For a slightly larger $r$ there are 3 connected components, and in the end there is one giant component. The dendrogram tracks how these connected components merge as the scale parameter increases: the result can be seen to the right of Fig. 16.
Persistent homology in dimension 0 will turn out to be closely related to that of a dendrogram provided that you forget precisely which points belong to each cluster. The idea is to consider a filtration of a simplicial complex $K$,

$$K_0 \subseteq K_1 \subseteq \ldots \subseteq K_n = K$$

and then apply simplicial homology to get a collection of vector spaces and linear maps

$$f_{i,j}^* : H_n(K_i) \to H_n(K_j)$$

for all $i \leq j$. The associated $n$-th persistent Betti numbers are given by $\beta_{i,j}^n = \dim \text{Im} f_{i,j}^*$.  

**Example 5.7.** Consider the filtration depicted in Fig. 17. Some persistent Betti numbers are:

- $\beta_{1,2}^0 = 1$,
- $\beta_{2,4}^0 = 2$,
- $\beta_{2,6}^2 = 2$,
- $\beta_{2,7}^2 = 1$,
- $\beta_{1,3}^2 = 1$, and $\beta_{1,10}^2 = 0$.  

Let $0 \neq [c] \in H_n(K_i)$. Then we say that $c$ is **born** at $K_i$ if $[c] \notin \text{Im} f_{i-1,i}^*$. If $[c]$ is born at $K_i$, then it **dies** at $K_j$ if it merges with an older class as we go from $K_{j-1}$ to $K_j$. That is, $f_{i,j-1}^*([c]) \notin \text{Im} f_{i-1,j-1}^*$ but $f_{i,j}^*([c]) \in \text{Im} f_{i-1,j}^*$.

**Example 5.8 (Example 5.7 continued).** We see that there is a class $[c] \in H_1(K_7)$ born at $K_7$ as well as a class $[c'] \in H_1(K_8)$ is born at $K_8$. These two classes ultimately coincide in $H_1(K_9)$. This means that the class $[c']$ born at $K_8$ vanishes at $K_9$ because it is merged with an older class. This is the **elder rule**: precedence is given to the class with the earliest birth time.

Define $\mu_{i,j}^n$ to be the number of classes in dimension $n$ born at index $i$ which die at index $j$. The number of classes born at index $i$ which never die is denoted by $\mu_{i,\infty}^n$.  

**Proposition 5.9.**

$$\mu_{i,j}^n = (\beta_{i,j}^n - \beta_{i,j}^{n-1}) - (\beta_{i-1,j}^{n-1} - \beta_{i-1,j}^{n-1,j})$$

**Proof.** The first term, $(\beta_{i,j}^{n-1} - \beta_{i,j}^{n})$, counts the number of classes born before or at index $i$ which die at index $j$. Accordingly, the second term, $\beta_{i-1,j}^{n-1} - \beta_{i-1,j}^{n-1,j}$, counts the number of classes born strictly before index $i$ which die at index $j$. Hence, $\mu_{i,j}^n$ is given by the difference of the two terms.  

\[ \square \]
The numbers $\mu_{n}^{i,j}$ are often visualized in either of the following ways:

- The **barcode in dimension** $n$ is obtained by interpreting every non-zero $\mu_{n}^{i,j}$ as the interval $[i, j)$ (representing the time when the "feature" was alive), and stacking all the intervals on top of each other in the plane; see Example 5.10.

- The **persistence diagram in dimension** $n$ is obtained by plotting $(i, j)$ (with multiplicity) in the plane for every non-zero $\mu_{n}^{i,j}$; see Example 5.10.

**Example 5.10** (Example 5.8 continued). We see that the only non-zero $\mu_{n}^{i,j}$’s are the following:

\[
\begin{align*}
\mu_{0}^{1,\infty} &= 1 & \mu_{0}^{2,7} &= 1 & \mu_{0}^{3,5} &= 1 & \mu_{0}^{4,6} &= 1 \\
\mu_{1}^{7,10} &= 1 & \mu_{1}^{8,9} &= 1 \\
\end{align*}
\]

The associated barcodes and persistence diagrams are shown in Fig. 18.
5.2.1 The Cech and Vietoris–Rips Filtrations

**The Cech Filtration**  For a given point set $P \subseteq \mathbb{R}^d$, we know from Theorem 4.10 that $\text{Cech}_r(P) \simeq \bigcup_{x \in P} B_r(x)$. In particular, given a filtration

$$\text{Cech}_r(P) \hookrightarrow \text{Cech}_{r_1}(P) \hookrightarrow \text{Cech}_{r_2}(P) \hookrightarrow \cdots \hookrightarrow \text{Cech}_{r_m}(P)$$

we get the following commutative diagram of vector spaces and linear morphisms

$$
\begin{array}{ccccccc}
H_n(P_{r_0}) & \rightarrow & H_n(P_{r_1}) & \rightarrow & H_n(P_{r_2}) & \rightarrow & \cdots & \rightarrow & H_n(P_{r_m}) \\
\downarrow \cong & & \downarrow \cong & & \downarrow \cong & & \downarrow \cong & & \downarrow \cong \\
H_n(\text{Cech}_{r_0}(P)) & \rightarrow & H_n(\text{Cech}_{r_1}(P)) & \rightarrow & H_n(\text{Cech}_{r_2}(P)) & \rightarrow & \cdots & \rightarrow & H_n(\text{Cech}_{r_m}(P))
\end{array}
$$

It is not difficult to prove that the persistent Betti numbers for the two rows coincide, and therefore the barcode/persistence diagram of the Cech filtration has a clear geometric interpretation: it captures the evolution of "holes" of a given dimension formed by balls of increasing radius around the points in $P$.

**The Vietoris-Rips Filtration**  For a given scale $r$, it need not be the case that $\text{VR}_r(P) \simeq P_r$ and therefore it may appear unclear how to interpret the persistent Betti numbers associated to a filtration of the form

$$\text{VR}_{r_0}(P) \hookrightarrow \text{VR}_{r_1}(P) \hookrightarrow \text{VR}_{r_2}(P) \hookrightarrow \cdots \hookrightarrow \text{VR}_{r_m}(P)$$

However, from Corollary 1.12 we have the following chain of inclusions

$$\text{Cech}_r(P) \subseteq \text{VR}_r(P) \subseteq \text{Cech}_d \delta r(P) \subseteq \text{VR}_{2 \delta r}(P)$$

where $\delta = \sqrt{\frac{d}{2(d+1)}}$, and in turn the following commutative diagram of vector spaces:

$$
\begin{array}{ccccccc}
H_n(\text{Cech}_r(P)) & \rightarrow & H_n(\text{Cech}_{2 \delta r}(P)) \\
\downarrow & & \downarrow \\
H_n(\text{VR}_r(P)) & \rightarrow & H_n(\text{VR}_{2 \delta r}(P))
\end{array}
$$

This diagram shows that any class in $H_n(\text{VR}_r(P))$ which has non-zero image in $H_n(\text{VR}_{2 \delta r}(P))$ must be non-zero in $H_n(\text{Cech}_{2 \delta r}(P))$. Importantly, this means that it represents a true topological feature of $P_{2 \delta r}(P)$. Conversely, any class of $H_n(\text{Cech}_r(P))$ which has non-zero image in $H_n(\text{Cech}_{2 \delta r}(P))$ must factor through $H_n(\text{VR}_r(P))$. Summarized: sufficiently long intervals in the barcode of the Cech filtration give rise to intervals in the barcode of the Vietoris–Rips filtration, and conversely, any sufficiently long interval in the barcode of the Vietoris–Rips filtration corresponds to a true geometrical feature. This relation will be made precise later in the course when we discuss the theory of interleavings.

With this obvious drawback of the Vietoris–Rips complex, one may wonder why consider it at all. There are primarily two reasons for doing so:

- Computation from a distance matrix. To construct the Cech complex one needs the coordinates of the points in $\mathbb{R}^d$, whereas the Vietoris–Rips is built exclusively from pairwise distances. Furthermore, this allows us to construct filtered simplicial complexes for points that do not embed in $\mathbb{R}^d$ for any $d$. An example could be a distance matrix obtained through correlations of genes.
Computations. Working with the Rips complex allows for certain heuristics to be implemented which in practice yield significant speed-ups.

**Example 5.11.** Any finite metric space has an associated Vietoris–Rips filtration. As an example, consider the metric space consisting of three points \{p_1, p_2, p_3, p_4\} and distance matrix

\[
\begin{bmatrix}
0 & 1 & 1 & 2 \\
1 & 0 & 2 & 1 \\
1 & 2 & 0 & 1 \\
2 & 1 & 1 & 0
\end{bmatrix}
\]

By defining \( \text{VR}_r(P) = \{ \sigma \subseteq P : \text{diam } \sigma \leq 2r \} \) as before, we obtain a filtration:

\[
\text{VR}_0(P) \subseteq \text{VR}_{1/2}(P) \subseteq \text{VR}_1(P).
\]

See Fig. 19 for a (non-isometric) visualization of this filtration.
6 Week 6: Algebraic Foundations

Last week we introduced the notion of persistent Betti numbers, and used them to compute the barcode. In this lecture we will give an algebraic characterization of the barcode. A more thorough discussion of the topics in this lecture can be found in any book on Quiver Representations.

6.1 Motivation

Consider the filtered simplicial complex $K$ shown in Fig. 20, together with the associated collection of vector spaces and linear maps

$$H_1(K_1) \to H_1(K_2) \to H_1(K_3) \to H_1(K_4) \to H_1(K_5).$$

From Proposition 5.9 it is an easy task to compute the barcode $B_1(K)$ of $K$ in dimension 1:

$$B_1(K) = \{[2,5], [3,4]\}.$$

This barcode clearly does not depend on the particular basis we choose at each index as the computation only amounts to rank considerations. From the algebraic point of view, however, there are some bases which are to be preferred. We will now illustrate this with two examples. In the first case, let $\{AB + BD + CD + AC\}, \{AB + BD + CD + AC, AC + CD + AD\}$ and $\{AB + BD + CD + AC\}$ be the bases for $K_2, K_3,$ and $K_4$. In the second case we replace $\{AB + BD + CD + AC, AC + CD + AD\}$ with $\{AB + BD + CD + AC, AB + BD + AD\}$. Representing the linear maps with respect to these bases yields:

$$
\begin{align*}
0 \to \mathbb{Z}_2 & \xrightarrow{[1]} \mathbb{Z}_2 \oplus \mathbb{Z}_2 \xrightarrow{[1,0]} \mathbb{Z}_2 \to 0. \\
0 \to \mathbb{Z}_2 & \xrightarrow{[1]} \mathbb{Z}_2 \oplus \mathbb{Z}_2 \xrightarrow{[1,1]} \mathbb{Z}_2 \to 0.
\end{align*}
$$

The former of these two splits into a direct sum:

$$(0 \to \mathbb{Z}_2 \xrightarrow{1} \mathbb{Z}_2 \xrightarrow{1} \mathbb{Z}_2 \to 0) \bigoplus (0 \to 0 \to \mathbb{Z}_2 \to 0 \to 0),$$

whereas the latter does not ”split up”. An advantage of such a decomposition is that the barcode can be easily read off from the support of its components (summands):

$$
\begin{align*}
\mu_{ij}^1 \left( (0 \to \mathbb{Z}_2 \xrightarrow{1} \mathbb{Z}_2 \xrightarrow{1} \mathbb{Z}_2 \to 0) \bigoplus (0 \to 0 \to \mathbb{Z}_2 \to 0 \to 0) \right) &= 1 \quad \text{if } (i,j) \in \{(2,5), (3,4)\} \\
\mu_{ij}^1 (0 \to \mathbb{Z}_2 \xrightarrow{1} \mathbb{Z}_2 \xrightarrow{1} \mathbb{Z}_2 \to 0) + \mu_{ij}^1 (0 \to 0 \to \mathbb{Z}_2 \to 0 \to 0) &= \begin{cases} 1 & \text{if } (i,j) \in \{(2,5), (3,4)\} \\ 0 & \text{otherwise}. \end{cases}
\end{align*}
$$
The goal of this section will be to prove that such a choice of basis always exists. This algebraic characterization of the barcode will be important when proving stability later in the course.

### 6.2 Persistence Modules

**Definition 6.1.** A partially ordered set (poset) \( P \) is a set \( P \) together with a binary relation \( \leq \) such that

1. \( p \leq p \), for all \( p \in P \),
2. \( p \leq q \) and \( q \leq p \) implies \( p = q \),
3. \( p \leq q \) and \( q \leq r \) implies \( p \leq r \).

If \( p \leq q \) and \( p \neq q \), then we write \( p < q \).

Here are some posets that appear in topological data analysis.

**Example 6.2.** The set \([n] = \{1, 2, \ldots, n\}\) is a partially ordered set under the standard ordering of the natural numbers. It will be convenient to represent this poset as a graph in the following way:

\[
\begin{array}{cccccc}
  \bullet & \rightarrow & 2 \bullet & \rightarrow & \cdots & \rightarrow & n-1 \bullet & \rightarrow & n \bullet
\end{array}
\]

More generally, \( \mathbb{N}, \mathbb{Z} \) and \( \mathbb{R} \) are posets in the obvious way.

**Example 6.3.** One can also assign a zigzag structure to \( \{1, 2, \ldots, n\} \):

\[
\begin{array}{cccccc}
  1 \bullet & \leftarrow & 2 \bullet & \rightarrow & \cdots & \leftarrow & n-1 \bullet & \rightarrow & n \bullet
\end{array}
\]

**Example 6.4.** The poset \( P = \{a, b, c, d\} \) with the binary relation \( a \leq b, c \leq d \):

\[
\begin{array}{cccc}
  b & \rightarrow & d \\
  \uparrow & & \uparrow \\
  a & \rightarrow & c
\end{array}
\]

**Definition 6.5.** A \( P \)-indexed persistence module (\( P \)-module) \( V \) is a collection of vector spaces \( \{V_p\}_{p \in P} \) and linear maps \( \{V(p \leq q) : V_p \rightarrow V_q\}_{p \leq q} \) such that

1. \( V(p \leq p) = \text{id} : V_p \rightarrow V_p \),
2. \( V(q \leq r) \circ V(p \leq q) = V(p \leq r) \).

**Remark 6.6.** A \( P \)-module is a functor from the partially ordered set \( P \) seen as a category in the obvious way, to the category of vector spaces.

**Example 6.7.** A filtered simplicial complex \( K_1 \subseteq K_2 \subseteq \cdots \subseteq K_n \) yields an \([n]\)-module for every integer \( i \),

\[
H_i(K_1) \rightarrow H_i(K_2) \rightarrow \cdots \rightarrow H_i(K_n).
\]

**Example 6.8.** We can also consider a zigzag of inclusions: \( K_1 \subseteq K_2 \supseteq K_3 \subseteq \cdots \supseteq K_n \),

\[
H_i(K_1) \rightarrow H_i(K_2) \leftarrow H_i(K_3) \rightarrow \cdots \leftarrow H_i(K_n).
\]
Example 6.9. Given a topological space $X$ and a real-valued function $f : X \to \mathbb{R}$, we get an $\mathbb{R}$-module $M$ by

$$M_t = H_n(f^{-1}(-\infty, t])$$

together with maps $M_t \to M_{t'}$ in homology induced by the inclusion $f^{-1}(-\infty, t] \subseteq f^{-1}(-\infty, t']$.

A morphism $f : V \to W$ between $P$-modules is a collection of linear maps $\{f_p : V_p \to W_p\}_{p \in P}$ such that the following diagram commutes for all $p \leq q$:

$$
\begin{array}{ccc}
V_p & \xrightarrow{V(p \leq q)} & V_q \\
\downarrow f_p & & \downarrow f_q \\
W_p & \xrightarrow{W(p \leq q)} & W_q.
\end{array}
$$

We say that $f$ is an epi-/mono-/isomorphism if $f_p$ is an epi-/mono-/isomorphism for all $p$.

Example 6.10. The following is an example of an epimorphism:

$$
\begin{array}{ccc}
V : & \mathbb{Z}_2 & \xrightarrow{1} & \mathbb{Z}_2 \\
\downarrow f & & \downarrow 1 & \downarrow 0 \\
W : & \mathbb{Z}_2 & \xrightarrow{0} & 0.
\end{array}
$$

Note that there are no non-zero morphisms from $W$ to $V$.

The direct sum of two $P$-modules $V$ and $W$ is the $P$-module $V \oplus W$ defined pointwise by $(V \oplus W)_p = V_p \oplus W_p$ and on linear maps by:

$$
\begin{bmatrix}
V(p \leq q) & 0 \\
0 & W(p \leq q)
\end{bmatrix}
\xrightarrow{V(p \leq q) \oplus W(p \leq q)} W_p \oplus W_q.
$$

Definition 6.11. We say that $V$ is decomposable if there exist non-zero $W, W'$ such that $V \cong W \oplus W'$. If no such $W, W'$ exist, then we say that $V$ is indecomposable.

For a $P$-module $V$, we define the total dimension of $V$ by $\dim V = \sum_{p \in P} \dim V_p$ whenever the sum exists.

Lemma 6.12. Let $P$ be finite and assume that $V_p$ is finite-dimensional for all $p$. Then

$$V \cong W^1 \oplus \cdots \oplus W^k$$

where each $W^i$ is indecomposable.

Proof. We will work inductively on the total dimension of $V$. If $\dim V = 1$ then $V$ is necessarily indecomposable. So assume that the statement holds for $\dim V \leq n$ and consider $V$ with $\dim V = n + 1$. If $V$ is indecomposable, then we are done. If not, we can write $V = W \oplus W'$ where

$$\dim V = \dim W + \dim W'$$

and both of the terms to the right are non-zero. Hence, by the inductive property, $W$ and $W'$ decompose into a direct sums of indecomposables. \qed
A submodule $V'$ of $V$ is a collection of subspaces $V'_p \subseteq V_p$ such that the following diagram commutes for all $p \leq q$

$$
\begin{cd}
V'_p & \longrightarrow & V'_q \\
\downarrow & & \downarrow \\
V_p & \longrightarrow & V_q
\end{cd}
$$

If there exists a non-trivial submodule $V'' \subseteq V$ such that $V = V' \oplus V''$, then we say that $V'$ and $V''$ are **summands** of $V$. For a morphism $f: V \rightarrow W$ we get submodules $\ker f \subseteq V$ and $\text{Im } f \subseteq W$ by taking pointwise kernels and images. These submodules are in general not summands.

**Example 6.13.**

$$
\begin{array}{ccc}
\ker(f) & 0 & \longrightarrow & 0 & \longrightarrow & \mathbb{Z}_2 \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
V & 0 & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & 1 & \mathbb{Z}_2 \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
W & \mathbb{Z}_2 & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & 0 \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\text{Im}(f) & 0 & \longrightarrow & \mathbb{Z}_2 & \longrightarrow & 0
\end{array}
$$

**Lemma 6.14 (Fitting Lemma).** Let $P$ be finite and $\dim V_p < \infty$ for all $p$. For every morphism $f: V \rightarrow V$ there exists an $n \geq 1$ such that $V = \text{Im } f^n \oplus \ker f^n$.

**Proof.** Since $f(V) \supseteq f^2(V) \supseteq f^3(V) \supseteq \cdots$ and $\dim V$ is finite, there must exists an $l$ such that $f^l(V) = f^{l+1}(V) = f^{l+2}(V) = \cdots$. Define $\phi := f^l: V \rightarrow V$. There exists for every $v \in V$ a $v' \in V$ such that $\phi(v) = \phi^2(v')$, and thus $v - \phi(v') \in \ker \phi$. This shows that $v = (v - \phi(v')) + \phi(v')$ where $v - \phi(v') \in \ker \phi$ and $\phi(v') \in \text{Im } \phi$. It remains to show that $\ker \phi \cap \text{Im } \phi = 0$.

If $v \in \ker \phi \cap \text{Im } \phi$, then $v = \phi(v')$ and $0 = \phi(v) = \phi^2(v')$. But $\phi: \phi(V) \rightarrow \phi^2(V)$ is an isomorphism, and therefore $v = \phi(v') = 0$. \hfill \Box

Now we restrict our attention to the poset $[n]$. Let $k$ be any field, and for $a < b \in [n] \cup \{\infty\}$, define $[a, b)$ to be the subset of $[n]$ given by the elements $\{i: a \leq i < b\}$. We define the **interval module** associated to $[a, b)$ to be the $[n]$-module $I^{[a,b)}$ defined by

$$
I^{[a,b)}_i = \begin{cases} k & \text{if } i \in [a, b) \\ 0 & \text{otherwise} \end{cases},
$$

together with the identity morphism $\text{id}: I^{[a,b)}_i \rightarrow I^{[a,b)}_j$ whenever $i, j \in [a, b)$.

**Example 6.15.** For $n = 2$ we have the following interval modules

$$
k \rightarrow 0 \quad 0 \rightarrow k \quad k \xrightarrow{\text{id}} k.
$$
Theorem 6.16. Let $V$ be an $[n]$-module such that $\dim V_p < \infty$ for all $p \in [n]$. Then

$$V \cong \bigoplus_{[a,b] \in B(V)} I^{[a,b]}$$

where $B(V)$ is a multiset of intervals in $[n]$ called the barcode of $V$.

Proof. We know from Lemma 6.12 that $V$ decomposes into a direct sum of indecomposables. Hence it suffices to show that if $V$ is indecomposable, then $V \cong I^{[a,b]}$ for some interval $[a,b]$. Assume without loss of generality that $V \neq 0$ and choose $0 \neq v \in V_1$. Let $b$ denote the smallest index for which $V(1 \leq b)(v) = 0$ and let $j = b - 1$. If no such $b$ exists, then let $b = \infty$ and $j = n$.

First we define a (well-defined) monomorphism $f : I^{[1,b]} \to V$ by $f_i(1_k) = V(1 \leq i)(v)$. Since $V_j$ is a vector space, we can extend $\{f_i(1_k)\}$ to a basis for $V_j$ and thus define an epimorphism $g_j : V_j \to k$ which satisfies $g_j(f_j(1_k)) = 1_k$. This extends to an (well-defined) epimorphism $g : V \to I^{[1,b]}$ by defining $g_i(w) = g_j(V(i \leq j)(w))$ for $i \leq j$, and 0 otherwise. Importantly,

$$g \circ f = \text{id}_{I^{[1,b]}}.$$

From Lemma 6.14 there exists an $m \geq 1$ such that

$$V = \text{Im}(f \circ g)^m \oplus \ker(f \circ g)^m.$$

Furthermore,

$$(f \circ g)^m(V) = f \circ (g \circ f) \circ (g \circ f) \circ \cdots \circ (g \circ f) \circ g(V) = f \circ g(V) = f(I^{[1,b]}),$$

where the last equality follows from $g$ being an epimorphism. Since $V$ was assumed to be indecomposable, and since $f$ is a monomorphism, it follows that

$$V = \text{Im}(f \circ g)^m = f(I^{[1,b]}) \cong I^{[1,b]}.$$

Proposition 6.17. Let $V$ be as in Theorem 6.16. Then $\mu_{i,j}^{[n]}$ counts the number of occurrences of $[i,j]$ in $B(V)$.

Proof. This follows from the following two observations

$$\mu_{[n]}^{i,j}(V) = \mu_{i,j}\left(\bigoplus_{[a,b] \in B(V)} I^{[a,b]}\right) = \sum_{[a,b] \in B(V)} \mu_{i,j}(I^{[a,b]}),$$

and

$$\mu_{i,j}(I^{[a,b]}) = \begin{cases} 1 & \text{if } (i,j) = (a,b) \\ 0 & \text{otherwise}. \end{cases}$$

Just as the dimension of a vector space is independent of the choice of basis, the previous result shows that although the particular decomposition in Theorem 6.16 is not unique, the associated multiset $B(V)$ is.

We state the following generalization of Theorem 6.16 without proof.
Theorem 6.18 (Crawley-Boevey (2015), Botnan & Crawley-Boevey (2020)). Let $T$ be a totally ordered set and $V$ a $T$-module such that $\dim V_t < \infty$ for all $t \in T$. Then

$$V \cong \bigoplus_{J \in B(V)} I^J$$

where $B(V)$ is a multiset of intervals in $T$.

An interval in $T$ is a subset $J$ with the property that if $x < y < z$ and $x, z \in J$ then $y \in J$. Interval modules are defined accordingly.
7 Week 7: The Persistence Algorithm

The algorithm for computing simplicial homology can be straightforwardly applied to compute persistent homology. In this lecture we will have a look at the algorithm and an improvement which makes computations faster in practice.

7.1 The standard algorithm

The standard algorithm for computing persistent homology is nothing more than the standard algorithm for simplicial homology with the ordering of the simplices given by the time they appear in the filtration. We illustrate this by computing the barcode of the filtration in Fig. 21. Expressing the boundary matrix in the order given by the barcode of the filtration in Fig. 21 yields:

\[
D = \begin{pmatrix}
1 & 2 & 3 & 4 & 13 & 12 & 24 & 34 & 14 & 134 & 124 \\
1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
4 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
24 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
34 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
134 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
124 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\end{pmatrix}
\]

and the corresponding identity matrix \( V = I_{12 \times 12} \). The resulting matrices \( R \) and \( V \) after
applying Algorithm 1 to $D$ are:

$$R = \begin{pmatrix}
1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
4 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\
12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
24 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
34 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
134 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
124 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\end{pmatrix}$$

and

$$V = \begin{pmatrix}
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
4 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
13 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
12 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
24 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
34 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
134 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
124 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
\end{pmatrix}$$

Observe that the simplicial homology of $K_i$ - the filtered simplicial complex at time $i$ - can be obtained by restricting $R$ and $V$ to the first $i$ columns.

Following the steps in Section 3.1.2:

$$\Sigma_{B_0} = \{(10100000000)^T, (11000000000)^T, (01010000000)^T \}$$

$$\Sigma_{Z_0} = \{(10000000000)^T, (01000000000)^T, (00100000000)^T, (00010000000)^T \}$$

$$\Sigma_{B_1} = \{(00001001100)^T, (00001111000)^T \}$$

$$\Sigma_{Z_1} = \{(000001111000)^T, (000001110100)^T \}.$$ 

The next step is to pair the vectors of $\Sigma_{Z_0}$ to those of $\Sigma_{B_0}$ based on their lowest entries:

$$(10000000000)^T \leftrightarrow \emptyset \quad (01000000000)^T \leftrightarrow (11000000000)^T$$

$$(00100000000)^T \leftrightarrow (10100000000)^T \quad (00010000000)^T \leftrightarrow (01010000000)^T$$

and in dimension 1:

$$(00001111000)^T \leftrightarrow (00001111000)^T \quad (000001110100)^T \leftrightarrow (000001001100)^T$$

From this we conclude that $H_0(K_{11}) \cong \mathbb{Z}_2$ is generated by the vertex $v_1$, and that $H_i(K_{11}) = 0$ for $i \geq 1$. However, from the point of view of persistent homology, we are interested in the
particular pairings: a pairing \( v \leftrightarrow w \) yields a bar \([t_v, t_w]\) in the barcode, where \( t_v \) is the time at which the cycle \( v \) appears in the filtration, and \( t_w \) is the time where \( w \) becomes a boundary. We see that \( t_v = \text{low}(v) \) and \( t_w = j \) where \( \text{low}(R_j) = t_v \). We summarize our findings in the following table:

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Birth</th>
<th>Death</th>
<th>Rep. cycle</th>
<th>Vertex notation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>( \infty )</td>
<td>((00000000000)^T)</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>2</td>
<td>6</td>
<td>((01000000000)^T)</td>
<td>2</td>
</tr>
<tr>
<td>0</td>
<td>3</td>
<td>5</td>
<td>((00100000000)^T)</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>4</td>
<td>7</td>
<td>((00010000000)^T)</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>8</td>
<td>11</td>
<td>((00001111000)^T)</td>
<td>13 + 12 + 24 + 34</td>
</tr>
<tr>
<td>1</td>
<td>9</td>
<td>10</td>
<td>((00001001100)^T)</td>
<td>13 + 34 + 14</td>
</tr>
</tbody>
</table>

This completes the computation of persistent homology. We summarize our findings:

1. Write down the boundary matrix \( D \) with the simplices ordered by their appearance in the filtration of a simplicial complex \( K \).
2. Use Algorithm 1 to obtain a decomposition \( R = DV \).
3. The columns \( R_i \) with \( \text{low}(i) \neq 0 \) form a basis \( \Sigma_B = \Sigma_{B_0} \cup \ldots \cup \Sigma_{B_d} \) for \( B_0(K) \oplus \ldots \oplus B_d(K) \).
4. The columns \( V_i \) such that \( R_i = 0 \) form a basis \( \tilde{\Sigma}_Z = \tilde{\Sigma}_{Z_0} \cup \ldots \cup \tilde{\Sigma}_{Z_d} \) for \( Z_0(K) \oplus \ldots \oplus Z_d(K) \).
5. For every \( n \)-cycle \( V_i \in \tilde{\Sigma}_Z \) do the following:
   (a) If there exists \( R_j \in \Sigma_B \) with \( \text{low}(R_j) = \text{low}(V_i) = i \), then output an interval \([i, j]\) in dimension \( n \).
   (b) If there exists no such \( R_j \), then output an interval \([i, \infty)\) in dimension \( n \).

At the algebraic level we also get bases at every filtration step:

1. Let \( \Sigma_{H_j} = \{ V_i \in \tilde{\Sigma}_Z : i \leq j \text{ and } \exists R_l \text{ where } j < l \text{ and } \text{low}(R_l) = i \} \).
2. Let \( \Sigma_{H_{j_1,j_2}} = \Sigma_{H_{j_1}} \cap \Sigma_{H_{j_2}} \).
3. Let \( \Sigma_{E_j} = \{ V_i \in \tilde{\Sigma}_Z \setminus \Sigma_{H_j} : i \leq j \} \).
4. It is a good exercise to verify the following properties
   (a) The set of \( n \)-cycles in \( \Sigma_{H_j} \cup \Sigma_{E_j} \) forms a basis for \( H_n(K_j) \).
   (b) The rank of the map \( H_n(K_{j_1}) \to H_n(K_{j_2}) \) is given by the cardinality of \( \Sigma_{H_{j_1,j_2}} \cup \Sigma_{E_{j_1}} \).

**Justification** Let us convince ourselves that our persistence algorithm outputs an interval \([i, j]\) if and only if \( \mu^{i,j} = 1 \). For every step of the filtration the addition of an \( n \)-simplex results in

\[
\dim H_n(K_j) = \dim H_n(K_{j-1}) + 1,
\]
or

\[
\dim H_{n-1}(K_j) = \dim H_{n-1}(K_{j-1}) - 1.
\]
The dimension increases if $R_j = 0$ (addition of a new cycle $V_j$) and decreases if $R_j \neq 0$ (addition of a boundary $R_j$). Assuming that the dimension decreases, then $\mu^{l,j}_{n-1} = 1$ for a unique $i < j$.

The persistence algorithm pairs $R_j$ to the unique vector $V_s \in \Sigma Z_{n-1}$ with low$(R_j) = \text{low}(V_s) = s$. Since $R_j$ is an $(n-1)$-boundary it can be written as a sum of $(n-1)$-cycles which satisfies the following property

$$R_j = V_s + \left( \text{sum of other vectors } V_k \in \Sigma Z_{n-1} \text{ with low}(V_k) < s \right).$$

This shows that $V_s$, which is generated at index $s$, gets identified with a sum of cycles appearing before index $s$ as we enter $K_j$. Consequently, $s = \text{low}(R_j)$, and $\mu^{s,j}_{n-1} = 1$. We conclude that the standard algorithm computes all the finite intervals correctly.

Now assume that $\mu^{\infty}_{n} = 1$. Since $\dim H_n(K_j) = \dim H_n(K_{j-1}) + 1$, we must have $\mu^{j,l}_{n} = 0$ for all $l < \infty$. If the persistence algorithm would pair $V_j$ to an $R_l$ for an $l > j$, then the preceding discussion shows that $\mu^{l,j}_{n} = 1$, contradicting that $\mu^{j,l}_{n} = 0$. Hence $V_j$ remains unpaired and the persistence algorithm also computes all infinite intervals correctly.

**Algebraic justification**  The standard algorithm actually provides us with enough data to decompose the homology persistence modules into summands isomorphic to interval modules. We saw above that the $n$-cycles of $\Sigma H_i \cup \Sigma E_i$ form a basis for $H_n(K_j)$. We modify the basis slightly: every $V_i \in \Sigma H_i$ is paired to an $R_l$ such that $R_l = V_i + \sum_k V_k$ where each $V_k$ in the sum appeared before $V_i$. We can thus replace the basis element $V_i$ with $R_l$, and then the morphism $H_n(K_{l-1}) \to H_n(K_l)$ maps the basis element $R_l$ of $H_n(K_{l-1})$ to 0. Hence, all the morphisms diagonalize and the persistence module decomposes into indecomposable summands. Let us illustrate this for the case $n = 1$. Using the basis as given by the persistence algorithm we get the bases (using vertex notation) \{12 + 13 + 24 + 34, 12 + 24 + 14\} and \{12 + 13 + 24 + 34\}, for $H_1(K_9)$ and $H_1(K_{10})$, respectively. However, the cycle 12 + 24 + 14 becomes equivalent to 12 + 13 + 24 + 14 in homology as we enter $K_{10}$, and therefore the linear maps take the form:

$$\text{span}\{12 + 13 + 24 + 34, 12 + 24 + 14\} \xrightarrow{[1\ 1]} \text{span}\{12 + 13 + 24 + 34\}.$$  

If we, however, replace 12 + 24 + 14 with the boundary to which it gets paired, then we obtain a new basis \{12 + 13 + 24 + 34, 13 + 34 + 14\} and the matrix diagonalizes to:

$$\text{span}\{12 + 13 + 24 + 34, 13 + 34 + 14\} \xrightarrow{[1 \ 0]} \text{span}\{12 + 13 + 24 + 34\}.$$  

And therefore $0 \to H_1(K_9) \to H_1(K_9) \to H_1(K_{10}) \to 0$ decomposes as

$$(0 \to \mathbb{Z}_2 \xrightarrow{1} \mathbb{Z}_2 \xrightarrow{1} \mathbb{Z}_2 \xrightarrow{0}) \bigoplus (0 \to 0 \to \mathbb{Z}_2 \to 0 \to 0),$$  

See the Section 6.1 for how this decomposition relates to the barcode.

We conclude that a pairing $V_i \leftrightarrow R_j$ corresponds to an interval summand with support $[i, j]$, and that an unpaired $V_i$ corresponds to an interval summand with support $[0, \infty)$.

**7.2 Clearing**

Consider the reduced matrix $R$ in the previous section. Since the column corresponding to the simplex 134 is non-zero, we know from the previous discussion that column number $i = \text{low}(134)$, which in this case is corresponds to the simplex 14, must be zeroed out. Likewise, we know that the column number $\text{low}(124)$, which corresponds to simplex 34, must be zeroed out. Hence, if we first reduced the 2-simplices, then we could skip the columns 14 and 34 when reducing the 1-simplices, as we know that they will be zeroed out. This is the whole idea behind clearing and can be summarized as follows:
• Let \( K \) be a filtered simplicial complex with \( \text{dim} \ K = d \).

• Use the standard algorithm to reduce the columns of \( D \) corresponding to \( d \)-simplices in order to obtain \( \Sigma_{B_{d-1}} \) and \( \Sigma_{Z_d} \).

• For every \( R_j \in \Sigma_{B_{d-1}} \), zero out column number \( \text{low}(R_j) \) in \( D \). Reduce the \((d-1)\)-simplices of the resulting boundary matrix to obtain the set \( \Sigma_{B_{d-2}} \). A basis \( \Sigma_{Z_{d-1}} \) for \( Z_{d-1}(K) \) is given by the set \( \Sigma_{B_{d-1}} \) together with all columns \( V_i \) corresponding to zero columns \( R_i \) that were not zeroed out prior to the reduction.

• Continue the same way until a bases for \( B_0(K) \) and \( Z_1(K) \) are computed.

The total number of rows that need not be reduced are

\[
\text{dim} \ B_{d-1}(K) + \text{dim} \ B_{d-2}(K) + \ldots + \text{dim} \ B_1(K).
\]

We do however need to reduce all columns corresponding to \( d \)-simplices, and in applications this number typically is much larger than the sum in Eq. (2). But the crux of the matter lies with the fact that reducing a column to zero is typically more expensive than reducing a column to become a boundary. And in the above case we still need to zero out a large number of columns corresponding to \( d \)-simplices. This turns out to be a major bottleneck and working with cohomology avoids this issue in practice.

Example 7.1. Let \( P \) be a set of 100 points in Euclidean space and say we want to compute the persistent homology of the Vietoris–Rips filtration in dimensions 0, 1 and 2. Counting the total number of 1, 2 and 3-simplices gives the total number of columns to be reduced in the naïve approach (0-simplices are already reduced):

\[
\binom{100}{2} + \binom{100}{3} + \binom{100}{4} = 4087875.
\]

Now (exercise)

\[
\text{dim} \ B_2(\text{VR}_\infty(P)) + \text{dim} \ B_1(\text{VR}_\infty(P)) = \binom{99}{3} + \binom{99}{2} = 161700.
\]

Clearing only marginally improves the computation time and the reason for this is the substantial number of 3-simplices that still need to be reduced in the initial step. Indeed, the number of columns which need to be zeroed out equals \( \text{dim} \ Z_3(\text{VR}_\infty(P)) = \binom{99}{4} = 3764376. \)
8 Week 9: Cohomology

Cohomology is a prominent tool in algebraic topology which carries more structure than homology, and homology and cohomology groups are often non-isomorphic. When computing homology and cohomology with coefficients in a field however, their vector spaces have the same dimensions. In particular, the barcodes of persistent homology and the barcodes of persistent cohomology of a filtered space are equivalent.

Somewhat surprisingly computing cohomology is often much faster in practice. The goal of this lecture is to introduce cohomology and see why its computation is faster.

8.1 Dual persistence modules

Recall that the dual of a vector space \( V \) over a field \( k \) is the vector space of linear maps \( V^* := \text{Hom}(V, k) \). It is well-known (and easy to prove) that if \( \{v_1, \ldots, v_d\} \) is a basis for \( V \), then \( \{\epsilon_1, \ldots, \epsilon_d\} \) is a basis for \( V^* \), where \( \epsilon_i(v_j) = 1 \) and \( \epsilon_i(v_j) = 0 \) for \( i \neq j \). Therefore \( \dim V = \dim V^* \) if \( V \) is finite-dimensional.

Furthermore, if \( f: V \rightarrow W \) is a linear map, then we get an induced linear map \( f^*: W^* \rightarrow V^* \) given by \( f^*(\alpha)(x) = \alpha(f(x)) \). If \( f \) is represented by the matrix \( A \), then \( f^* \) is represented in the dual bases by the matrix \( A^T \). It is clear that \( (g \circ f)^* = f^* \circ g^* \).

Consider the interval module \( I_{[a,b)} \) for an interval \( [a,b) \) in \( [m] \):

\[
0 \rightarrow \cdots \rightarrow 0 \rightarrow k_1 \rightarrow k_1 \rightarrow \cdots \rightarrow k_1 \rightarrow k \rightarrow 0 \rightarrow \cdots \rightarrow 0.
\]

Its dual \( (I_{[a,b)})^* \) is isomorphic to the interval module \( I_{[m+1-b,m+1-a]} \) where \( (m+1-b, m+1-a] \) is the interval \( \{i \in [m] : m+1-b < i \leq m+1-a\} \).

**Example 8.1.** The dual of \( 0 \rightarrow k_1 \rightarrow k_1 \rightarrow k \) is isomorphic to \( k_1 \rightarrow k_1 \rightarrow k \rightarrow 0 \).

The following proposition follows from the previous two observations and Theorem 6.16

**Proposition 8.2.** Let \( V \) be an \([m]\)-module such that \( \dim V_i < \infty \) for all \( i \in [m] \). Then \( V^* \cong \bigoplus_{[a,b) \in B(V)} I_{[m+1-b,m+1-a]} \).

In particular, \( B(V^*) = \{(m+1-b, m+1-a] : [a,b) \in B(V)\} \).

\[3\text{This is in fact an if and only if statement but the converse is non-trivial to prove.}\]
8.2 Simplicial Cohomology

In the following we are working over the field $k = \mathbb{Z}_2$.

**Definition 8.3.** Let $n \geq 0$ be an integer and $K$ a simplicial complex. The vector space of $n$-cochains in $K$ is the dual vector space $C^n(K) = C_n(K)^* = \text{Hom}(C_n(K), \mathbb{Z}_2)$.

There is a coboundary operator $\partial^n : C^n(K) \to C^{n+1}(K)$ defined on $\alpha : C_n(K) \to \mathbb{Z}_2$ by

$$\partial^n(\alpha)(c) = \alpha(\partial_{n+1}(c))$$

where $\partial_n$ is the boundary operator in homology. If $\partial^n(\alpha) = 0$, then we say that $\alpha$ is an $n$-cocycle, and if $\alpha = \partial^{n-1}(\beta)$, then we say that $\alpha$ is an $n$-coboundary. The $n$-cocycles and $n$-coboundaries define vector spaces $Z^n(K)$ and $B^n(K)$, respectively.

We immediately observe that

$$\partial^{n+1} \circ \partial^n(\alpha)(c) = \alpha(\partial_{n+1} \circ \partial_{n+2}(c)) = \alpha(0) = 0.$$  \hspace{1cm} (3)

It follows that $B^n(K) \subseteq Z^n(K)$ as in homology.

**Definition 8.4.** The $n$-th simplicial cohomology vector space of a simplicial complex $K$ is the quotient vector space $H^n(K) = Z^n(K)/B^n(K)$.

For every $n$-simplex $\sigma$ in $K$ we have a dual $\sigma^* : C_n(K) \to \mathbb{Z}_2$ given by

$$\sigma^*(\tau) = \begin{cases} 1 & \text{if } \sigma = \tau, \\ 0 & \text{otherwise}. \end{cases}$$

By representing the coboundary operator in all dimensions in the bases given by the duals of simplices, we can apply Algorithm 1 to compute $H^n(K)$ in all dimensions. For the moment we shall order the columns and rows in the opposite order of what we would do in homology. The reason for doing this will become clear when we turn to persistent cohomology.

**Example 8.5.** Consider the simplicial complex in Fig. 22. For $\sigma = \{3, 4\} \in K$ we get $\partial^1(\sigma^*) = \tau = \{2, 3, 4\}^*$, and for $v = \{2\}$ we get $\partial^0(v^*) = \{2, 3\}^* + \{2, 4\}^* + \{1, 2\}^*$. Representing the coboundary operator as mentioned above yields,

$$M = \begin{pmatrix} 234 & 34 & 23 & 24 & 13 & 12 & 4 & 3 & 2 & 1 \\ 234 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 34 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 23 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 24 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 13 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

46
Applying Algorithm 1 to the above matrix yields the following $R$ and $V$, respectively:

\[
\begin{pmatrix}
234 & 34 & 23 & 24 & 13 & 12 & 4 & 3 & 2 & 1 \\
234 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
34 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
23 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\
24 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
13 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
12 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\end{pmatrix}
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\end{pmatrix}
\]

Using vector notation as in Section 3.1.2, we find bases for $Z^2(K), B^2(K), Z^1(K), B^1(K), Z^0(K)$:

\[
\begin{align*}
\tilde{\Sigma}_{Z^2} & = \{(1000000000)^T\} \\
\Sigma_{B^2} & = \{(1000000000)^T\} \\
\tilde{\Sigma}_{Z^1} & = \{(0110000000)^T, (0101000000)^T, (0000100000)^T, (0000010000)^T\} \\
\Sigma_{B^1} & = \{(0101000000)^T, (0110100000)^T, (0011010000)^T\} \\
\tilde{\Sigma}_{Z^0} & = \{(0000001111)^T\}.
\end{align*}
\]

By means of Lemma 2.20 we extend $\Sigma_{B^i}$ to a basis $\Sigma_{Z^i}$ for $Z^i(K)$ as in the homology case:

\[
\Sigma_{Z^2} = \Sigma_{B^2}, \quad \Sigma_{Z^1} = \{(0110000000)^T\} \cup \Sigma_{B^1}, \quad \Sigma_{Z^0} = \{(0000001111)^T\}.
\]

We conclude that $H^2(K) = 0$, $H^1(K) \cong \mathbb{Z}_2$ and $H^0(K) \cong \mathbb{Z}_2$.

The fact that $H^n(K) \cong H_n(K)$ can easily be proven by showing that $\dim H^n(K) = \dim Z^n(K) - \dim B^n(K) = \dim Z_n(K) - \dim B_n(K) = \dim H_n(K)$. We shall give an alternative proof. First we need the following characterization of cycles and cocycles.

**Lemma 8.6.**

1. $\gamma \in Z^n(K)$ if and only if $\gamma(c) = 0$ for all $c \in B_n(K)$,
2. $\gamma \in B^n(K)$ if and only if $\gamma(c) = 0$ for all $c \in Z_n(K)$.

**Proof.**
1. Assume $\gamma \in Z^n(K)$ and $c = \partial_{n+1}(d)$. Then $\gamma(c) = \gamma(\partial_{n+1}(d)) = \partial^n(\gamma)(d) = 0$. Conversely, if $\gamma$ is not in $Z^n(K)$, then there must exist some $d$ such that $0 \neq \partial^n(\gamma)(d) = \gamma(\partial_n(d))$.

2. Assume $\gamma(c) = 0$ for all $c \in Z_n(K)$. Define $\beta : B_{n-1}(K) \to \mathbb{Z}_2$ by $\beta(d) = \gamma(c)$ where $d = \partial_n(c)$. This is well-defined, because if $d = \partial_n(c) = \partial_n(c')$, then $\partial_n(c - c') = 0$ and thus $c - c' \in Z_n(K)$. It follows that $\gamma(c) = \gamma(c')$. By extending a basis for $B_{n-1}(K)$ to a basis for $C_{n-1}(K)$, we can extend $\beta$ to a linear map $\beta : C_{n-1}(K) \to \mathbb{Z}_2$ satisfying

$$\partial^{n-1}(\beta)(c) = \beta(\partial_n(c)) = \gamma(c).$$

We conclude that $\gamma = \partial^{n-1}(\beta)$. The converse statement is straightforward.

**Lemma 8.7.** The morphism $h : H^n(K) \to H_n(K)^*$ given by $h([\alpha])\langle c \rangle = \alpha(c)$ is well-defined, and it is an isomorphism.

**Proof.** Let $\alpha = \alpha' + \partial^n(\beta)$. Then

$$h([\alpha])\langle c \rangle = \alpha(c) = \alpha'(c) + \partial^n(\beta)(c) = \alpha'(c) + \beta(\partial_n(c)) = \alpha'(c)$$

where the last equality follows from $c$ being an $n$-cycle. This shows that $h$ is independent of the representative cocycle. Next, let $c = c' + \partial_{n+1}(d)$,

$$\alpha(c) = \alpha(c') + \alpha(\partial_{n+1}(d)) = \alpha(c') + \partial^n(\alpha)(d) = \alpha(c'),$$

since $\alpha$ is an $n$-cocycle. This shows that $h$ is well-defined.

- $h$ is injective: If $h([\alpha]) = 0$, then $\alpha(c) = 0$ for all $c \in Z_n(K)$. It follows from Lemma 8.6 that $\alpha \in B^n(K)$, hence $[\alpha] = 0$.

- $h$ is surjective: Let $f : H_n(K) \to \mathbb{Z}_2$. Pre-composing $f$ with the linear map $Z_n(K) \to H_n(K)$, defines a $f' : Z_n(K) \to \mathbb{Z}_2$. By extending a basis for $Z_n(K)$ to a basis for $C_n(K)$, we can lift $f'$ to a linear map $f' : C_n(K) \to \mathbb{Z}_2$. Since $f'(b) = f([b]) = f(0) = 0$ for $b \in B_n(K)$, it follows that $f' \in Z^n(K)$ by Lemma 8.6. Hence $h([f']) = f$.

Since $\dim V = \dim V^*$ for $V$ finite-dimensional, it follows that

$$H_n(K) \cong H_n(K)^* \cong H^n(K).$$

### 8.3 Persistent Cohomology

For an inclusion $f : K_i \subseteq K_j$ of simplicial complexes, there is an induced map $f^* : H^n(K_j) \to H^n(K_i)$ given by

$$f^*([\alpha])(c) = \alpha(f(c)).$$

We leave it as an exercise to verify that this map is well-defined.

To a filtration $K_1 \subseteq K_2 \subseteq \cdots \subseteq K_m$ we thus get the following **cohomology persistence module**:

$$H^n(K) : \quad H^n(K_m) \to H^n(K_{m-1}) \to \cdots \to H^n(K_2) \to H^n(K_1).$$

Let $H_n(K)$ denote the associated homology persistence module,

$$H_n(K) : \quad H_n(K_1) \to H_2(K) \to \cdots \to H_n(K_{m-1}) \to H_n(K_m).$$
Lemma 8.8. The persistence modules $H^n(K)$ and $H_n(K)^*$ are isomorphic.

Proof. For $i \in [m]$, define $h_i : H^n(K_i) \to H_n(K_i)^*$ by $h_i(\alpha)(c) = \alpha(c)$. That yields the following commutative diagram for all $i \leq j$

\[
\begin{array}{ccc}
H^n(K_j) & \longrightarrow & H^n(K_i) \\
\downarrow h_j & & \downarrow h_i \\
H_n(K_j)^* & \longrightarrow & H_n(K_i)^*
\end{array}
\]

where the vertical morphisms are isomorphisms by Lemma 8.7.

The following theorem is an immediate consequence of the previous lemma and Proposition 8.2.

Theorem 8.9. There is a bijection between the barcodes $B(H^n(K))$ and $B(H_n(K))$ given by

$B(H_n(K)) \ni [a,b) \leftrightarrow (m+1-b,m+1-a] \in B(H^n(K))$.

Remark 8.10. In fact, if we view $H^n(K)$ as an persistence module indexed over $[m]$ with the opposite ordering,

$H^n(K_1) \leftarrow H^n(K_2) \leftarrow \cdots H^n(K_{m-1}) \leftarrow H^n(K_m),$

then the barcode of $H^n(K)$ coincides with the barcode of $H_n(K)$.

8.3.1 Computation

It is tempting to think that persistent cohomology can be computed by applying the standard algorithm to the coboundary matrix as in Example 8.5. This turns out to almost be the case.

Example 8.11. Returning to Example 8.5 and pairing vectors based on their lowest non-zero entry as in the standard algorithm for persistent homology we get the pairs

\[
\begin{align*}
(1000000000)^T & \leftrightarrow (1000000000)^T \\
(0101000000)^T & \leftrightarrow (0101000000)^T \\
(0000100000)^T & \leftrightarrow (0110100000)^T \\
(0110000000)^T & \leftrightarrow \emptyset \\
(0000001111)^T & \leftrightarrow \emptyset
\end{align*}
\]

Let us start with the first pairing, which is between $\{2,3,4\}$ and the column corresponding to $\{3,4\}^*$. Since the duals of simplices appear in the opposite order of the simplices, this results in an interval $[1,2)$ in dimension 2 (no, this is not a typo). Similarly we get intervals $[4,7),[5,8), [6,9)$ and $[3,\infty)$ in dimension 1. Lastly we get an interval $[10,\infty)$ in dimension 0.

The barcode of the previous example can easily be transformed into the barcode of persistent cohomology, and thus persistent homology. One can prove that applying the standard algorithm to the boundary matrix as above computes the barcode of the following persistence module in relative cohomology:

$H^n(K,K_{m-1}) \to H^n(K,K_{m-2}) \to \cdots \to H^n(K,K_1) \to H^n(K)$.

The barcode of relative cohomology corresponds to the barcode of $H^n(K)$ according to the following table:
The reader may easily verify that transforming the barcode of Example 8.11 into the homology barcode yields the expected result.

### 8.3.2 Clearing

The salient point of the last subsection is that the persistent homology barcode can be computed by means of cohomology. A na"ive implementation will not improve computational speed much compared to the standard algorithm. However, the speed improves dramatically if we employ the clearing technique discussed in the previous lecture.

Just as in the homology case, we know that if the column corresponding to the \( n \)-cochain \( \sigma^* \) forms an \( n \)-coboundary with lowest non-zero entry at row \( i \), then the column corresponding to the \( (n+1) \)-cochain added at time \( i \) will be reduced to 0 under Algorithm 1. This means that if we have reduced all the \( (n-1) \)-cochains to form a basis for \( B_n(K) \), then we avoid reducing \(|B^n(K)|\) columns when computing bases for \( Z^n(K) \) and \( B^{n+1}(K) \). This is the whole idea behind clearing and can be summarized as follows:

- Let \( K \) be a filtered simplicial complex with dim \( K = d \).
- Let \( D \) be the coboundary operator in matrix form, with the columns ordered in the opposite order of the filtration (as in Example 8.5).
- Use the standard algorithm to reduce the columns of \( D \) corresponding to 0-simplices in order to obtain \( \Sigma B_1 \) and \( \hat{\Sigma} Z_0 \).
- For every \( R_j \in \Sigma B_1 \), zero out column number low\((R_j)\) in \( D \). Reduce the 1-cochains of the resulting boundary matrix to obtain the set \( \Sigma B_2 \). A basis \( \hat{\Sigma} Z_1 \) for \( Z^1(K) \) is given by the set \( \Sigma B_1 \) together with all columns \( V_i \) corresponding to zero columns \( R_i \) that were not zeroed out prior to the reduction.
- Continue the same way until a bases for \( Z^{d-1}(K) \) and \( B^d(K) \) are computed.

The total number of columns that need not be reduced are

\[
\dim B^1(K) + \dim B^2(K) + \ldots + \dim B^{d-1}(K).
\] (4)

The difference with homology is that we avoid zeroing out a large number of columns corresponding to \( d \)-simplices. Let us return to Example 7.1 in the case of cohomology.

**Example 8.12.** When working with cohomology we know that \( Z^3(K) = C^3(K) \) if restricted to a simplicial complex without 4-simplices. Hence, if we were to na"ively reduce the coboundary matrix without using clearing we would have to do a total number of

\[
\dim C^0(\text{VR}_\infty(P)) + \dim C^1(\text{VR}_\infty(P)) + \dim C^2(\text{VR}_\infty(P)) = 100 + 100 + 100 = 300.
\]

column reductions. This number is much lower than for the na"ive homology algorithm, but the computations are not faster in practice because the columns to be reduced are much larger. Furthermore, the total number of columns to be zeroed out equals

\[
\dim Z^2(\text{VR}_\infty(P)) + \dim Z^1(\text{VR}_\infty(P)) + \dim Z^0(\text{VR}_\infty(P)) = 99 + 99 + 1 = 199.
\]
With clearing on the other hand, we need not zero out the following number of columns:

$$\dim B^1(\text{VR}_\infty(P)) + \dim B^2(\text{VR}_\infty(P)) = \binom{99}{1} + \binom{99}{2} = 4950.$$  

In conclusion: 161800 column reductions but only a single column will be reduced to 0.

The previous example illustrates one of the reasons why cohomology with clearing typically performs much faster than homology with clearing: the number of columns to zero out becomes small, and it is comparatively much cheaper to reduce a column to a boundary. Furthermore, it turns out that in practice a large number of the boundary columns need no reduction. More on this can be found in [1].
9 Week 10: Stability of Persistent Homology

In this lecture we shall discuss metrics on barcodes/persistence diagrams and discuss to what extent persistent homology is stable under perturbation of the input data.

9.1 The Bottleneck Distance

Let \( C \) and \( D \) be multisets of intervals \( \langle a, b \rangle \) in \( \mathbb{R} \). Here the notation \( \langle a, b \rangle \) denotes that the interval can be any well-defined member of \{\([a, b], [a, b), (a, b], (a, b)\)\}. We define a matching between \( C \) and \( D \) to be a collection of pairs \( \chi = \{(I, J) \in C \times D\} \) where each \( I \) and each \( J \) can occur in at most one pair. Equivalently, a matching is a bijection between a subset of \( C \) and a subset of \( D \). If \( (I, J) \in \chi \), then we say that \( I \) is matched to \( J \). If \( I \) does not appear in a pair, then we say that \( I \) is unmatched.

Example 9.1. Let \( C = \{I_1, I_2\} \) and \( D = \{J\} \). Then the pair \( (I_1, J) \) defines a matching where \( I_2 \) is unmatched. Note that \( \{(I_1, J), (I_2, J)\} \) does not define a matching.

Define the cost \( c(I, J) \) of matching \( I = \langle a, b \rangle \) to \( J = \langle c, d \rangle \) to be

\[
c(I, J) = \max(|c - a|, |d - b|),
\]

and define the cost \( c(I) \) of not matching \( I \) to be

\[
c(I) = (b - a)/2.
\]

Given this we define the cost of a matching \( \chi \) to be

\[
c(\chi) := \max \left( \sup_{(I, J) \in \chi} c(I, J), \sup_{I \in C \cup D} c(I) \right).
\]

We say that \( \chi \) is an \( \varepsilon \)-matching if \( c(\chi) \leq \varepsilon \).

Remark 9.2. A geometric interpretation of this is as follows: consider the intervals of \( C \) and \( D \) as points in \( \mathbb{R}^2 \). An \( \varepsilon \)-matching then pairs up points \( p \) and \( q \), corresponding to intervals of \( C \) and \( D \) respectively, such that \( ||p - q||_\infty \leq \varepsilon \), and such that any unmatched point is at most \( \varepsilon \) away from the diagonal in the \( l_\infty \)-norm.

Lemma 9.3. If there exists an \( \varepsilon \)-matching between \( C \) and \( D \), and an \( \varepsilon' \)-matching between \( D \) and \( E \), then the compositions of matching yields an \( (\varepsilon + \varepsilon') \)-matching between \( C \) and \( E \).

Proof. Let \( I = \langle a_1, b_1 \rangle \in C \), \( J = \langle a_2, b_2 \rangle \in D \) and \( K = \langle a_3, b_3 \rangle \in E \) such that \( I \) is matched to \( J \), and \( J \) is matched to \( K \). Then \( |a_3 - a_1| \leq |a_3 - a_2| + |a_2 - a_1| \leq \varepsilon + \varepsilon' \), and similarly for \( |b_3 - b_1| \).

If \( I \) and \( J \) are as above, but \( J \) is unmatched in the second matching, then by assumption we have \( b_2 \leq a_2 + 2\varepsilon' \). It follows that

\[
a_1 + 2\varepsilon + 2\varepsilon' \geq a_2 + \varepsilon + 2\varepsilon' \geq b_2 + \varepsilon \geq b_1,
\]

and thus \( c(I) \leq \varepsilon + \varepsilon' \).

A similar argument applies if \( J \) is matched to \( K \) under the second matching, but unmatched under the first matching.

It follows from the definition of a matching and the previous lemma that the following defines a distance.
Figure 23: Left: A graph (black) and its perturbation (red). Right: The persistence diagrams of the 0-th sublevel set persistent homology of the two graphs.

**Definition 9.4.** The **bottleneck distance** between $C$ and $D$ is

$$d_B(C, D) = \inf \{c(\chi) : \chi \text{ is a matching between } C \text{ and } D\}.$$  

**Example 9.5.** Let $C = \{[0, 6), [4, 6)\}$ and $D = \{[2, 8), [10, 12)\}$. The trivial matching which leaves everything unmatched has a cost of the half the length of the longest interval: $(8 - 2)/2 = 3$. Hence, $d_B(C, D) \leq 3$. In order to improve this bound we need to match the interval $[2, 8)$ to either $[0, 6)$ or $[4, 6)$. These two pairings come at the same expense

$$c([0, 6), [2, 8)) = c([4, 6), [2, 8)) = 2.$$  

If we match $[0, 6)$ with $[2, 8)$, then we can either match $[4, 6)$ to $[10, 12)$ or leave both of the unmatched. We see that leaving them both unmatched comes at the lowest cost of

$$\max\{c([0, 6), [2, 8)), c([4, 6)), c([10, 12))\} = 2.$$  

If we however match $[4, 6)$ to $[2, 8)$, then the best we can do is

$$\max\{c([4, 6), [2, 8)), c([0, 6)), c([10, 12))\} = 3.$$  

We conclude that $d_B(C, D) = 2$.

The previous example was somewhat tedious but luckily there are fast algorithms for computing the bottleneck distance in practice. By recasting the problem as an instance of a *bipartite matching problem*, state-of-the-art algorithms compute $d_B(C, D)$ in $O(n^{1.5} \log n)$ where $n = |C| + |D|$.  

**9.1.1 Stability of Persistent Homology**

One of the key properties of persistent homology is the fact that it is **stable** in the bottleneck distance. As an example, consider the two graphs shown to the left in Fig. 23. Associated to these graphs we get persistence diagrams in dimension 0 by considering sublevel sets and applying homology in dimension 0. The persistence diagram of the perturbed graph has two points close to the points of the original persistence diagrams, together with two points close to the diagonal (noise). This suggests a matching where the latter two points (intervals) are left unmatched, and the two other points (intervals) will be matched with a cost bounded by the perturbation in the $l_\infty$-norm.
More generally, let $X$ be a topological space equipped with two functions $f, g: X \to \mathbb{R}$. Associated to these functions we have $\mathbb{R}$-modules $M^f$ and $M^g$ given by

$$M^f_t = H_i(f^{-1}(-\infty, t]) \quad \quad M^g_t = H_i(g^{-1}(-\infty, t]).$$

(6)

Assuming that $\dim M^f_t < \infty$ and $\dim M^g_t < \infty$ for all $t$, which ensures that we have well-defined barcodes by Theorem 6.18, the following is true

**Theorem 9.6.**

$$d_B(B(M^f), B(M^g)) \leq ||f - g||_\infty.$$ 

We shall prove this theorem for sufficiently *tame* functions.

Similar stability results apply to point sets: Fig. 24 shows a noisy sample from the circle and its perturbation, together with the persistence diagrams in dimension 1 of the associated to the Vietoris–Rips filtrations. For a finite point set $P$, let $H_i(\text{VR}(P))$ denote the $\mathbb{R}$-indexed persistence module defined at $t$ by $H_i(\text{VR}_t(P))$ and for which the morphism $H_i(\text{VR}_s(P)) \to H_i(\text{VR}_t(P))$ is induced by the inclusion $\text{VR}_s(P) \subseteq \text{VR}_t(P)$.

**Theorem 9.7.** Let $P$ and $Q$ be finite sets of points in $\mathbb{R}^d$, and $\sigma: P \to Q$ a bijection such that $||p - \sigma(p)|| \leq \epsilon$ for all $p \in P$. Then,

$$d_B(B(H_i(\text{VR}(P))), B(H_i(\text{VR}(Q)))) \leq \epsilon.$$

By means of the *Gromov–Hausdorff distance* this theorem can be strengthened to include point sets that need not be embedded in a common space nor have the same cardinality. We will not concern ourselves with the more general result.

A stability theorem for the Čech filtration in the case that $P$ and $Q$ need not have the same cardinality follows from Theorem 9.6. Let $H_i(\text{Čech}(P))$ denote the $\mathbb{R}$-module associated to the Čech filtration. Recall the notation $P_r := d_\infty^{-1}(-\infty, r]$ introduced in Section 5.1.

**Theorem 9.8.** Let $P$ and $Q$ be finite sets of points in $\mathbb{R}^d$. If there exists an $\epsilon \geq 0$ such that $P \subseteq Q_\epsilon$ and $Q \subseteq P_\epsilon$, then

$$d_B(B(H_i(\text{Čech}(P))), B(H_i(\text{Čech}(Q)))) \leq \epsilon.$$

**Proof.** Define functions $d_P, d_Q : \mathbb{R}^d \to \mathbb{R}$ by

$$d_P(x) = \min_{p \in P} ||x - p|| \quad \quad d_Q(x) = \min_{q \in Q} ||x - q||.$$

Fix an $x \in \mathbb{R}^2$ and let $p \in P$ be a point closest to $x$. Then, from the assumption that $P \subseteq Q_\epsilon$, there exists some $q \in Q$ with $||p - q|| \leq \epsilon$. This shows that $d_P(x) \leq d_Q(x) + \epsilon$, and symmetrically, $d_Q(x) \leq d_P(x) + \epsilon$. We conclude that $||d_P - d_Q||_\infty \leq \epsilon$.

Now define $\mathbb{R}$-modules $M^{d_P}$ and $M^{d_Q}$ as above. From Theorem 4.10 we know that $M^{d_P} \cong H_i(\text{Čech}(P))$ and $M^{d_Q} \cong H_i(\text{Čech}(Q))$. Since isomorphic $\mathbb{R}$-modules necessarily have the same barcode, we conclude that

$$d_B(B(H_i(\text{Čech}(P))), B(H_i(\text{Čech}(Q)))) = d_B(B(M^{d_P}), B(M^{d_Q})) \leq ||d_P - d_Q||_\infty \leq \epsilon.$$
9.1.2 Generalizations

The bottleneck distance is only sensitive to the maximum over a matching. There are other "more discriminative" distances which are more frequently used in data analysis, although they do not exhibit the same type of stability as the bottleneck distance. One such distance is the degree $q$ Wasserstein distance,

$$W_q(C, D) = \inf_{\text{matchings } \chi \text{ between } C \text{ and } D} \left( \sum_{(I, J) \in \chi} c(I, J)^q + \sum_{\text{unmatched } I \in C \cup D} c(I)^q \right)^{1/q},$$

where $q$ is some positive real number. The bottleneck distance can be obtained as a limit in the following way:

$$\lim_{q \to \infty} W_q(C, D) = d_B(C, D).$$

9.2 Interleavings

We will now introduce the theory of interleavings and an associated interleaving distance between persistence modules. Importantly, and as we will see, the interleaving distance coincides with the bottleneck distance, and it is easy to show stability in the interleaving distance. Hence, our stability results will follow from this equivalence.

**Discrete setting** For the moment, let us assume that we are working with persistence module indexed by the integers. A 0-interleaving is nothing more than an isomorphism, i.e. a collection of morphisms $\{M_i \to N_i\}$ and a collection of morphisms $\{N_i \to M_i\}$ such that the following diagrams commute:
A 1-interleaving is a collection of slanted morphisms such that the following diagram commutes:

\[
\begin{array}{cccccc}
\cdots & \rightarrow & M_i & \rightarrow & M_{i+1} & \rightarrow & M_{i+2} & \rightarrow & \cdots \\
\uparrow & & \downarrow & & \downarrow & & \uparrow & & \downarrow \\
\cdots & \rightarrow & N_i & \rightarrow & N_{i+1} & \rightarrow & N_{i+2} & \rightarrow & \cdots \\
\end{array}
\]

Note that a 1-interleaving requires that the morphism \( M_i \rightarrow M_{i+2} \) factors through \( N_{i+1} \) in a natural way. Intuitively, this means that every “2-persistent” feature of \( M_i \)—a bar of length at least two that overlaps the index \( i \)—is also present in \( N_i \), and vice versa. A 2-interleaving is a collection of morphisms \( \{ M_i \rightarrow N_{i+2} \} \) and \( \{ N_i \rightarrow M_{i+2} \} \) making a similar diagram commute.

It should now be possible to deduce what an \( n \)-interleaving is.

**The formal definition**  Now we shall assume that the modules \( M \) and \( N \) are indexed over the real numbers. Define the \( \epsilon \)-shift of \( M \) to be the persistence module \( M^\epsilon \) defined by \( M^\epsilon_t = M_{t+\epsilon} \) and \( M^\epsilon(s \leq t) = M(s + \epsilon \leq t + \epsilon) \) for all \( s \leq t \in \mathbb{R} \). If \( f : M \rightarrow N \) is a morphism, then we get an \( \epsilon \)-shifted morphism \( f^\epsilon : M^\epsilon \rightarrow N^\epsilon \) defined by \( f^\epsilon_t = f_{t+\epsilon} \).

**Example 9.9.** The \( \epsilon \) shift of the interval module \( I^[a,b) \) is \( I^[a-\epsilon,b-\epsilon) \).

Let \( \eta^\epsilon_M : M \rightarrow M^\epsilon \) be the morphism whose restriction to each \( M_t \) is the internal morphism \( M_{t+\epsilon} \).

**Definition 9.10.** Given \( \epsilon \in [0, \infty) \), an \( \epsilon \)-interleaving between \( M \) and \( N \) is a pair of morphisms \( \psi : M \rightarrow N^\epsilon \) and \( \varphi : N \rightarrow M^\epsilon \) such that \( \varphi^\epsilon \circ \psi = \eta^\epsilon_M \) and \( \psi^\epsilon \circ \varphi = \eta^\epsilon_N \). We say that \( M \) and \( N \) are \( \epsilon \)-interleaved.

Explicitly, the last condition states the following two diagrams commute for all \( t \in \mathbb{R} \):

\[
\begin{array}{ccc}
M_t & \xrightarrow{\psi_t} & N_{t+\epsilon} \\
\downarrow & & \downarrow \\
M_{t+\epsilon} & \xrightarrow{\varphi_t} & N_{t+2\epsilon} \\
\end{array}
\]

\[
\begin{array}{ccc}
N_t & \xrightarrow{\phi_t} & M_{t+\epsilon} \\
\downarrow & & \downarrow \\
N_{t+\epsilon} & \xrightarrow{\psi_{t+\epsilon}} & M_{t+2\epsilon} \\
\end{array}
\]

**Lemma 9.11.** If \( M \) and \( N \) are \( \epsilon \)-interleaved, and \( N \) and \( L \) are \( \epsilon' \)-interleaved, then \( M \) and \( L \) are \( (\epsilon + \epsilon') \)-interleaved.

**Proof.** Let \( \psi : M \rightarrow N^\epsilon \), \( \varphi : N \rightarrow M^\epsilon \), \( \psi' : N \rightarrow L^\epsilon' \) and \( \varphi' : L \rightarrow N^\epsilon' \) be the respective morphisms satisfying the interleaving condition. They yield a morphism \( \psi'' : M \rightarrow L^{\epsilon + \epsilon'} \) through the composition

\[
M_t \xrightarrow{\psi_t} N_{t+\epsilon} \xrightarrow{\psi'_{t+\epsilon}} L_{t+\epsilon + \epsilon'},
\]

and a morphism \( \varphi'' \circ M \rightarrow L^{\epsilon + \epsilon'} \) given by

\[
L_t \xrightarrow{\varphi'_{t+\epsilon}} N_{t+\epsilon} \xrightarrow{\varphi_{t+\epsilon + \epsilon'}} M_{t+\epsilon + \epsilon'}.
\]
It remains to check that \((\varphi''_{t} + \epsilon') \circ \psi''_{t} = \eta_{M}^{2(\epsilon + \epsilon')}\) and \((\psi''_{t} + \epsilon') \circ \varphi''_{t} = \eta_{L}^{2(\epsilon + \epsilon')}\). The first equality follows from the following commutative diagram:

\[
\begin{array}{c}
M_{t} \xrightarrow{M_{(t \leq t + 2\epsilon)}} M_{t + 2\epsilon} \\
\downarrow \varphi_{t+\epsilon} \downarrow \\
N_{t + \epsilon} \xrightarrow{N_{(t + \epsilon \leq t + 2\epsilon')}} N_{t + \epsilon + 2\epsilon'} \\
\downarrow \psi_{t+\epsilon} \downarrow \\
L_{t + \epsilon + \epsilon'} \xrightarrow{\psi_{t+\epsilon+\epsilon'}} L_{t + \epsilon + \epsilon'}
\end{array}
\]

The second equality follows from a similar commutative diagram.

It follows from the definition of an interleaving and the previous lemma that the following indeed defines a metric.

**Definition 9.12.** The *interleaving distance* between two \(\mathbb{R}\)-modules \(M\) and \(N\) is

\[
d_{I}(M, N) = \inf \{ \epsilon : \text{there exists an } \epsilon\text{-interleaving between } M \text{ and } N \}.\]

The following result is fundamental to topological data analysis and non-trivial. We defer its proof to a later lecture.

**Theorem 9.13.** Let \(M\) and \(N\) be \(\mathbb{R}\)-modules such that \(M_{t}\) and \(N_{t}\) are finite-dimensional for all \(t \in \mathbb{R}\). Then,

\[
d_{B}(B(M), B(N)) = d_{I}(M, N).
\]

An immediate corollary is a proof of Theorem 9.6

**Proof of Theorem 9.6.** Since \(\|f - g\|_{\infty} \leq \epsilon\), we have the following commutative diagram of inclusions for every \(t \in \mathbb{R}\):

\[
\begin{array}{c}
f^{-1}(-\infty, t) \xrightarrow{f^{-1}(-\infty, t + \epsilon)} f^{-1}(-\infty, t + 2\epsilon) \\
\downarrow \downarrow \\
g^{-1}(-\infty, t) \xrightarrow{g^{-1}(-\infty, t + \epsilon)} g^{-1}(-\infty, t + 2\epsilon)
\end{array}
\]

Applying \(H_{t}\) yields the following commutative diagram

\[
\begin{array}{c}
M_{t}^f \xrightarrow{M_{t+\epsilon}^f} M_{t+2\epsilon}^f \\
\downarrow \downarrow \\
M_{t}^g \xrightarrow{M_{t+\epsilon}^g} M_{t+2\epsilon}^g
\end{array}
\]

It is easy to see that this defines an \(\epsilon\)-interleaving pair.

**Proof of Theorem 9.7.** This proof is analogous to the previous proof. Label the points of \(P = \{p_1, \ldots, p_m\}\) and \(Q = \{q_1, \ldots, q_m\}\) such that \(\|p_i - q_i\| \leq \epsilon\). If \(\sigma = \{p_{i_1}, \ldots, p_{i_m}\}\) is a simplex of \(\text{VR}_{r}(P)\), then we have by definition of the Vietoris–Rips complex that \(\text{diam } \sigma \leq 2r\). From the bijection we get that the corresponding set \(\tau = \{q_{i_1}, \ldots, q_{i_m}\}\) satisfies \(\text{diam } \tau \leq 2r + 2\epsilon\). Hence,
we have an inclusion \( \text{VR}_r(P) \subseteq \text{VR}_{r+\epsilon}(Q) \) defined by the simplicial mapping which sends the vertex \( p_i \) to \( q_i \). Symmetrically we have an inclusion \( \text{VR}_r(Q) \subseteq \text{VR}_{r+\epsilon}(P) \). Putting these two relations together we find

\[
\begin{align*}
\text{VR}_r(P) & \subseteq \text{VR}_{r+\epsilon}(Q) \subseteq \text{VR}_{r+2\epsilon}(P) \\
\text{VR}_r(Q) & \subseteq \text{VR}_{r+\epsilon}(P) \subseteq \text{VR}_{r+2\epsilon}(Q)
\end{align*}
\]

Applying \( H_i \) to the above relations defines an \( \epsilon \)-interleaving between \( H_i(\text{VR}(P)) \) and \( H_i(\text{VR}(Q)) \). It follows from Theorem 9.13 that

\[
d_B(B(H_i(\text{VR}(P))), B(H_i(\text{VR}(Q)))) = d_I(H_i(\text{VR}(P)), H_i(\text{VR}(Q))) \leq \epsilon.
\]

\[\square\]
10 Week 11: Clustering and TDA

Clustering is the task of partitioning data into (disjoint) subsets, such that data belonging to different partitions differ among themselves more than data which belong to the same partition. Put topologically, clustering is the task of inferring the connected components of the data. Fig. 25 shows that there is a wide range of available clustering methods. In this section we shall focus on methods inspired by topology.

10.1 Some Examples and Kleinberg’s Theorem

We briefly list some classical clustering methods.

The $\epsilon$-neighborhood graph Given a finite set of points $P$ and a distance $d(p_i, p_j)$ between every pair of points $p_i, p_j \in P$, we define the $\epsilon$-neighborhood graph to be the graph on the vertex set $P$, with an edge between $p_i$ and $p_j$ if $d(p_i, p_j) \leq \epsilon$. In other words, the 1-skeleton of $\text{VR}_\epsilon(P)$. The set of points $P$ is then partitioned according to the following rule: $p_i$ and $p_j$ belong to the same partition if and only if they belong to the same connected component of the neighborhood graph.

$kNN - k$-th nearest neighbor graph An alternative to the $\epsilon$-neighborhood is to define a graph by drawing an edge between the point $p_i$ and $p_j$ if $p_j$ is amongst the $k$ closest points of $p_i$.

$k$-means clustering Assuming that $P$ is a finite set of points in $\mathbb{R}^d$ and $k > 0$ is an integer, the goal of $k$-means clustering is to partition the objects into $k$ clusters such that intra-cluster
If \( f \) we must have that

\[
\sum_{i=1}^{k} \sum_{p \in P_i} ||p - \mu_i||^2
\]

is minimal. Here \( \mu_i \) denotes the mean of the points in \( P_i \). Computing such an optimal partition is an NP-hard problem, and iterative techniques are employed to find local minima. This technique is ill-suited if your clusters are non-convex.

**DBSCAN** A density based refinement of the \( \epsilon \)-neighborhood graph which is suited for non-convex domains is as follows: let \( m \) be a positive integer and let \( C \) denote the set of points in \( P \) which have at least \( m \) points within distance \( \epsilon \). The elements of \( C \) are called the core points. Now, construct the \( \epsilon \)-neighborhood graph on the core points, and use the \( \epsilon \)-neighborhood graph approach to cluster the points of \( C \). Any other point \( p \in P \) which is not a core point, is then added to the partition containing the core point \( p_i \) if \( d(p_i, p) \leq \epsilon \). If \( p \) is not within \( \epsilon \) distance to any core point, then \( p \) is labelled as noise.

**Kleinberg’s Theorem** Let \( P \) be a finite set with \(|P| \geq 2 \). A clustering function (on \( P \)) is a function which takes as input a distance function \( d: P \times P \rightarrow \mathbb{R}_{\geq 0} \) and outputs a partition \( \Gamma \) of \( P \). We say that a clustering function \( f \) is isometry invariant if \( f(d) = f(d') \) whenever \((P, d)\) and \((P, d')\) are isometric. If \( f(\alpha \cdot d) = f(d) \) for all \( \alpha > 0 \), then \( f \) is scale invariant. Lastly, let \( d' \) be such that \( d'(p_i, p_j) \leq d(p_i, p_j) \) whenever \( p_i, p_j \) are clustered together in \( f(d) \), and \( d'(p_i, p_j) \geq d(p_i, p_j) \) whenever \( p_i \) and \( p_j \) do not belong to the same partition. If \( f(d) = f(d') \) for all such \( d \) and \( d' \), then we say that \( f \) is consistent.

**Theorem 10.1** (Kleinberg). If \( f \) is isometry-invariant, scale-invariant and consistent, then

1. (lump) \( f(d) = \{P\} \) for all distances \( d \), or
2. (discrete) \( f(d) = \{\{p\} : p \in P\} \) for all distances \( d \).

Proof. Let \(|P| = n \) and let \( d \) be such that \( d(p_i, p_j) = 1 \) for all \( p_i \neq p_j \). By isometry invariance we must have that \( f(d) = \{P\} \) or \( f(d) = \{\{p\} : p \in P\} \). Assume the former is the case, and let \( d' \) be any other distance on \( P \), and rescale such that \( \alpha d' < d \). Then

\[
f(d') \leq_{\text{scale inv}} f(\alpha d') \leq_{\text{consistency}} f(d) = \{P\}.
\]

If \( f(d) = \{\{p\} : p \in P\} \), then scale such that \( \alpha d' > d \) and apply a similar argument.

An issue with clustering is that there might be no unique correct scale at which the data should be considered. Indeed, Fig. [26] illustrates that what appears to be well-defined clusters at one scale, may reveal a finer structure upon inspection at a smaller scale. One may attempt to rectify these issues by considering **hierarchical clustering methods**. Such methods do not assign a single partition to the input metric space, but rather a one-parameter family of clusters capturing the evolution of the clusters as the connectivity parameter increases - much like persistent homology. The output of such a method is typically a graph like the one shown in Fig. [27]. These methods do however also have their deficiencies but we will not discuss this here.
Figure 26: Example illustrating the multiscale nature of the clustering problem. At a coarse scale, two clusters (dark shading) are apparent. At a finer scale, each of these two clusters appear to decompose further into two subclusters (light shading).

Figure 27: Dendrogram for the point set shown in Fig. 26.

10.2 ToMATo

ToMATo - Topological Mode Analysis Tool - is the result of combing the classical hill climbing algorithm with persistent homology. The algorithm first appeared in [5] and in this section we follow their exposition.

To illustrate the idea of ToMATo in the continuous setting, consider the graph shown to the left in Fig. 28 which depicts an example density-function of an assumed underlying distribution from which we sample. As there are two clear maxima, we say that each of them represent a cluster, and the points belonging to the respective clusters are those points belonging to their respective ascending regions. That is, the ascending region of a critical point \( m \) is the set of all points \( x \) in the domain which end up at \( m \) by following the flow induced by the gradient vector field of the density function. In the example in the question, this means that every point to the left of the valley where the two peaks meet are clustered to one cluster, and all the points to the right of the valley are clustered to the other cluster. But what if we consider the density function to the right in Fig. 28 - defining every local peak to be a cluster seems counter-productive. To proceed one chooses a threshold \( \tau > 0 \) and defines a local maxima to define a cluster if the associated peak has a prominence of at least \( \tau \). That is, one needs to descend a height of at least \( \tau \) in order to reach a higher peak. Persistent homology is in turn used to select the appropriate \( \tau \): let \( f : X \to \mathbb{R} \) denote the density function and consider the persistent homology of the superlevel filtration given by

\[
H_0(f^{-1}[s, \infty)) \to H_0(f^{-1}[t, \infty))
\]

for \( s \geq t \). In Fig. 29 we show the right-most function of Fig. 28 together with its persistence diagram. Note that the diagram is drawn upside down as we are considering the superlevel
filtration and not the sublevel filtration we have frequently encountered. The blue line in the figure is the separating line which is used to distinguish the noisy peaks from the peaks that are to define clusters. Note that the line defines the parameter $\tau$ by means of the the equation $y = x - \tau$. Once we have settled on such a $\tau$ we select the prominent peaks $m_p$ and for a peak $m$ of prominence less than $\tau$, we consider the minimum $\epsilon$ such that $m$ belongs to the same connected component as a $\tau$-prominent peak $m_p$ in $f^{-1}[m - \epsilon, \infty)$. Any point which belongs to the ascending region of $m$ will now be associated to the point $m_p$, and the clusters are defined accordingly. E.g., in Fig. 29 the points to left of the deepest valley define a cluster, and the points to the right define another cluster.

10.2.1 The Discrete Case

For the purpose of this lecture we shall assume that we are given a point sample $P$ and an estimate $\hat{f}(p)$ of the density at every point $p \in P$. Such an estimate can be obtained in numerous ways, the most naïve being a simple normalized count of the number of points of $P$ within a $\delta$-ball around $p$ for some $\delta > 0$.

Fix an $\epsilon > 0$ and consider the $\epsilon$-neighborhood graph $G$ constructed from the points in $P$. With this information we can form our initial clusters: every local maximum, i.e. all points $p_i$
Figure 2: Our approach in a nutshell: (a) estimation of the underlying density function \( f \) at the data points; (b) result of the basic graph-based hill-climbing step; (c) approximate PD showing 2 points far off the diagonal corresponding to the 2 prominent peaks of \( f \); (d) final result obtained after merging the clusters of non-prominent peaks.

Figure 30: The pipeline of the discrete algorithm (copied from [5]).

satisfying \( \hat{f}(p_i) > \hat{f}(p_j) \) for all neighbors of \( p_i \) in \( G \), represents a cluster. If \( p_j \) is not a local maximum then we associate \( p_j \) to the same cluster as its neighbor with the greatest filtration value. If there are \( n \) local maxima, then we get a collection of clusters \( \{C_1, \ldots, C_n\} \).

By defining the filtration value \( \hat{f}(e) \) of an edge \( e \) connecting \( p_i \) and \( p_j \) to be \( \max(\hat{f}(p_i), \hat{f}(p_j)) \) we obtain a superlevel filtration of the graph \( G \). The algorithm proceeds precisely as in the continuous case: compute the persistence diagram in dimension 0 of the superlevel filtration of \( \hat{f} : G \rightarrow \mathbb{R} \) and use the persistence diagram to determine a threshold parameter \( \tau \). Upon choosing a \( \tau \), iterate over the points in the order given by \( \hat{f}(p_1) > \hat{f}(p_2) > \ldots > \hat{f}(p_n) \) and merge clusters as follows:

- If \( \hat{f}(p_i) \) is a local maximum in \( G \), then do nothing.
- If \( \hat{f}(p_i) \) is not a local maximum, then \( p_i \) belongs to a cluster \( C_k \). Iterating (in an arbitrary order) over the neighbors \( p_j \) of \( p_i \) in \( G \) with \( \hat{f}(p_j) \geq \hat{f}(p_i) \), we merge the cluster \( C_l \) containing \( p_j \) with \( C_k \) if

  \[
  \min(\max\{\hat{f}(p) : p \in C_l\}, \max\{\hat{f}(p) : p \in C_k\}) < \hat{f}(p_i) + \tau
  \]

  That is, we replace \( C_k \) and \( C_l \) in \( C \) by \( C_l \cup C_k \).

The resulting \( C \) contains the clusters. The four steps of the discrete algorithm and the end result is shown in Fig. 30. This illustration is copied from [5] and the reader is encouraged to have a look at that paper for other nice examples. Note that the above explanation was written in a conceptual way and that an implementation would iterate over the points using a union-find structure. We include the algorithm from [5] in Fig. 31.
The mode-seeking phase takes a linear time in the size of $G$ once the vertices have been sorted. As for the merging phase, it makes $O(n)$ union and $O(m)$ find queries to the union-find data structure $U$, where $n$ and $m$ are respectively the number of vertices and the number of edges of $G$. If an appropriate representation is used for $U$ (e.g. a disjoint-set forest [16]), and if the vertex gradients and the entry roots are stored in separate containers with constant-time access (e.g. arrays), then the worst-case time complexity of Algorithm 1 becomes $O(n \log n + m \alpha(n))$, where $\alpha$ stands for the inverse Ackermann function.

As for the space complexity, note that the graph $G$ does not have to be stored entirely in main memory, since only the neighborhood of the current vertex $i$ is involved at the $i$-th iteration of the clustering procedure. The main memory usage is thus reduced to $O(n)$, where $n$ is the number of vertices of $G$. The total space complexity remains $O(n + m)$ though, as the graph needs to be stored somewhere (e.g. on the disk).

**ALGORITHM 1:** Clustering

*Input:* simple graph $G$ with $n$ vertices, $n$-dimensional vector $\tilde{f}$, real parameter $\tau \geq 0.$

Sort the vertex indices $\{1, 2, \cdots, n\}$ so that $\tilde{f}(1) \geq \tilde{f}(2) \geq \cdots \geq \tilde{f}(n)$;

Initialize a union-find data structure $U$ and two vectors $g, r$ of size $n$;

for $i = 1$ to $n$ do

\hspace{1em} Let $N$ be the set of neighbors of $i$ in $G$ that have indices lower than $i$;

\hspace{1em} if $N = \emptyset$ then

\hspace{2em} // vertex $i$ is a peak of $\tilde{f}$ within $G$

\hspace{2em} Create a new entry $e$ in $U$ and attach vertex $i$ to it;

\hspace{2em} $r(e) \leftarrow i;$ \quad // $r(e)$ stores the root vertex associated with the entry $e$

\hspace{1em} else

\hspace{2em} // vertex $i$ is not a peak of $\tilde{f}$ within $G$

\hspace{2em} $g(i) \leftarrow \arg\max_{j \in N} \tilde{f}(j);$ \quad // $g(i)$ stores the approximate gradient at vertex $i$

\hspace{2em} $e_i \leftarrow U.f\text{ind}(g(i));$

\hspace{2em} Attach vertex $i$ to the entry $e_i$;

\hspace{2em} for $j \in N$ do

\hspace{3em} $e \leftarrow U.f\text{ind}(j);$ \quad // $e$ stores the entry of $j$

\hspace{3em} if $e \neq e_i$ and $\min\{\tilde{f}(r(e)), \tilde{f}(r(e_i))\} < \tilde{f}(i) + \tau$ then

\hspace{4em} $U.\text{union}(e, e_i);$

\hspace{4em} $r(e \cup e_i) \leftarrow \arg\max_{\{r(e), r(e_i)\}} \tilde{f};$

\hspace{4em} $e_i \leftarrow e \cup e_i;$

\hspace{3em} end

\hspace{2em} end

\hspace{1em} end

*Output:* the collection of entries $e$ of $U$ such that $\tilde{f}(r(e)) \geq \tau$.

Figure 31: The algorithm copied from the journal version of [5].
Figure 32: The mapper pipeline for a dense sample of the circle together with the height function.

10.3 A First Taste of Mapper

Mapper was first introduced in [12]. We now give a very brief introduction to a simplified version of Mapper and postpone a more thorough discussion until next week.

The idea is to start with a finite metric space $P$ together with a function $f: P \to \mathbb{R}$. One may choose $f$ as one likes, and typical examples include density and eccentricity estimates. Furthermore, we shall assume that $\mathbb{R}$ is covered by a collection of intervals $I_1, I_2, \ldots, I_k$ such that $I_i \cap I_{i+1} \neq \emptyset$ for all $i < k$ and such that no three intervals have a common intersection.

The algorithm now proceeds as follows:

1. Let $G$ denote the empty graph.

2. For every interval $I_j$ cluster the points of $f^{-1}(I_j)$ according to some pre-defined clustering algorithm. A typical choice is to consider the $\epsilon$-neighborhood graph of $f^{-1}(I_j)$ for some fixed $\epsilon > 0$.

3. For every cluster in $f^{-1}(I_j)$ add a vertex to $G$.

4. We draw an edge between any two vertices in $G$ if their respective clusters have a non-empty intersection.

**Example 10.2.** A simple example is shown in Fig. 32. Here we assume a dense sampling of the circle, and the function is the height function. Observe that covering the real line with two covering elements would result in a graph without a loop.

A more interesting example taken from [12] is shown in Fig. 33.
11 Week 12: Reeb Graphs and Mapper Continued

The Mapper construction introduced last week can be seen as an approximate Reeb graph.

11.1 Reeb Graphs

Let $X$ be a topological space and $f: X \to \mathbb{R}$ a continuous function. We define the Reeb graph associated to the pair $(X, f)$ to be the quotient space $R(X, f) = X/\sim$ where $\sim$ is the equivalence relation generated by $x \sim y$ if $f(x) = f(y) = t$ and $x$ and $y$ belong to the same path component of the fiber $f^{-1}(t)$. Note that this space in general is not a graph but it will be for reasonable pairs such as 1) $M$ is a compact manifold and $f$ is a generic smooth function, and 2) $X$ is a compact polyhedron and $f$ is a piecewise-linear map. By the following lemma from undergraduate topology we see that $f$ induces a well-defined continuous map $\hat{f}: R(X, f) \to \mathbb{R}$ given by $\hat{f}([x]) = f(x)$.

Lemma 11.1. Let $X/\sim$ be the quotient of a topological space $X$ by an equivalence relation $\sim$. Any continuous function $f: X \to Y$ which is constant on equivalence classes induces a continuous function $X/\sim \to Y$ with respect to the quotient topology.

Reeb graphs were originally introduced in the context of Morse theory to study the evolution level sets of real-valued functions.

Example 11.2. In Fig. 34 we see a typical example of a Reeb graph and how the graph captures connectivity information as well as "loops" in the original manifold. This illustrates a lemma from [8] that says that the Reeb graph associated to a closed orientable 2-manifold of genus $g$ has exactly $g$ loops.

Clearly the Reeb graph loses a lot of higher-dimensional connectivity information of the space $X$. It is however always true that the surjection $X \to X/\sim$ induces an isomorphism $H_0(X) \cong H_0(X/\sim)$ and an epimorphism $H_1(X) \to H_1(X/\sim)$. In particular, if $X$ is connected and simply-connected, then $R(X, f)$ must be a tree regardless of the function $f$.

By replacing $\mathbb{R}$ with $\mathbb{R}^d$ we can associate a Reeb space to a function $f: X \to \mathbb{R}^d$ precisely as we did for the Reeb graph. The resulting spaces are typically quite complicated.

11.2 The Continuous Mapper

An open cover of a topological space $X$ is a collection $\mathcal{U} = \{U_i\}_{i \in I}$ of open sets such that $\bigcup_{i \in I} U_i = X$. Associated to a finite cover there is a simplicial complex $N(\mathcal{U})$ called the nerve of $\mathcal{U}$ containing simplices $\{\sigma \subseteq I: \cap_{i \in \sigma} U_i \neq \emptyset\}$. Assuming a continuous function $f: X \to \mathbb{R}^d$, and a finite open cover $\mathcal{U}$ of $\mathbb{R}^d$, we get a pullback cover of $X$ defined by

$$f^*\mathcal{U} = \{f^{-1}(U_i) : U_i \in \mathcal{U}\}.$$
We can further refine the pullback cover $f^*U$ by splitting every open set $f^{-1}(U_i)$ into its path components. We denote the resulting refined cover by $\hat{f}^*U$. The (continuous) mapper associated to the pair $(X, f)$ is the simplicial complex $M(X, f) := N(\hat{f}^*U)$.

**Example 11.3.** Fig. 35 shows a continuous map on a surface and the resulting Mapper graph.

While the resulting simplicial complex of the previous example turned out to be a graph, that need not be the case in general. We do however have the following result:

**Lemma 11.4.** The dimension of $M(X, f)$ is bounded by the dimension of $N(U)$.

**Proof.** If $V$ and $V'$ are different path-connected components of some $f^{-1}(U_i)$ then clearly $V \cap V' = \emptyset$. Let $V_i$ be a path-component of $f^{-1}(U_i)$ and assume that $\bigcap_{i=1}^{m+1} V_i$ is non-empty. By the previous assumption we must have that each $U_i$ is distinct, and furthermore

$$\emptyset \neq f(\bigcap_{i=1}^{m+1} V_i) \subseteq \bigcap_{i=1}^{m+1} U_i.$$  

We conclude that if $M(X, f)$ contains an $m$-simplex, then so does $N(U)$. 

We define the **resolution** of a cover to be $res(U) = \sup \{ \text{diam } U_i : U_i \in U \}$. One can think mapper $M(X, f)$ as an approximation to $R(X, f)$ as the resolution of the cover goes to 0 (this can be made precise): rather than probing the connectivity information over every fiber, we only consider the connectivity information over preimages of (small) open sets. It is no surprise that as long as the fibers are ”locally constant” that the Mapper graph should resemble the Reeb graph for covers of low enough resolution. The next example illustrates this.
Example 11.5. Consider the Reeb graph and two mapper graphs shown in Fig. 36. The first cover is rather coarse and fails to detect any of the two loops, whereas the cover to the right captures the larger loop but an even smaller resolution would be needed to detect both loops.

11.3 Mapper in Practice

Given a finite set of data $P$ together with a metric (or dissimilarity measure) between data points, the user is typically left to decide the filter function $f: P \to \mathbb{R}^d$ and the cover of $\mathbb{R}^d$. The outputted Mapper graph (simplicial complex) will be highly sensitive in this parameters. Furthermore, in order to compute path components (clusters) of pre-images, there is typically yet another parameter involved in the construction of the Mapper graph. Assuming for the moment that $f$ and $U$ are provided, and that we have a way of clustering data, then the Mapper algorithm is as follows:

**Data:** A data set $P$ with a pairwise distance measure $d$, a function $f: P \to \mathbb{R}^d$ and a cover $U$ of $\mathbb{R}^d$.

**Result:** The mapper graph associated to $(P, d)$.

Let $K = \emptyset$.

for $U \in U$ do
  Decompose $f^{-1}(U)$ into clusters $C_{U,1}, \ldots, C_{U,k_U}$.
  Add a vertex $v_{U,i}$ to $K$ for every cluster $C_{U,i}$.
end

if $C_{U,1,j_1} \cap \cdots \cap C_{U,m,j_m}$ then
  Add the simplex $\{v_{U,1,j_1}, \ldots, v_{U,m,j_m}\}$ to $K$.
end

return $K$.

**Algorithm 2:** The Mapper algorithm.

Some remarks on the parameters in question:

**The filter function $f$** Examples of typical functions encountered in the literature.

- For data in $\mathbb{R}^d$ one can project down onto a lower-dimensional subspace, e.g. by means of principal component analysis or by a non-linear dimension reduction algorithm.
Figure 36: Left: the Reeb graph of a pair \((X, f)\). Middle: the Mapper graph for a coarse cover. Right: The Mapper graph for a cover of smaller resolution.

Figure 37: Mapper from a finite data set.
• Eccentricity measures, in an intuitive sense, how far from the "center" of the data a given point is. That is,

\[ E_r(p) = \left( \frac{\sum_{q \in P} d(p, q)^r}{N} \right)^{1/r} \]

where \( 1 \leq r \leq \infty \).

• Density estimates such as Gaussian kernel,

\[ \delta_\epsilon(p) = \sum_{q \in P} \exp\left(-\frac{d(p, q)^2}{\epsilon}\right), \]

where \( \epsilon > 0 \).

**The cover \( \mathcal{U} \)** For covers of \( \mathbb{R} \) it is customary to let \( \mathcal{U} \) consist of regularly spaced intervals of equal resolution, and such that neighbouring cover elements overlap with a percentage below 50\%. For covers of \( \mathbb{R}^d \) one typically considers hypercubes in the analogue way. The resulting Mapper will in general not be graphs but for ease of visualization it is customary to merely output the 1-skeleton. Note that the output of the algorithm is very sensitive to the chosen cover.

**Clustering** The most common choices are \( kNN \) clustering, DBSCAN and a simple \( \epsilon \)-neighborhood graph. Note that one can approach the clustering problem in two ways: one can either cluster all the data immediately and then computing connected components of preimages can be done by means of standard graph algorithms. Another, more local, approach is to cluster each preimage separately. A local computation may be a better option when dealing with very large data.
12 Week 13: Generalized Persistence

It is now time to look at homology groups indexed by more general posets than the linear one.

12.1 Zigzag Persistent Homology

A zigzag poset on $n$ vertices is a partial order of the form

$$
1 \leftrightarrow 2 \leftrightarrow \cdots \leftrightarrow n-1 \leftrightarrow n
$$

where $\leftrightarrow$ denotes that the arrow can be either $\leftarrow$ or $\rightarrow$. For instance, these are both examples of zigzag posets on $n$ vertices:

$$
\begin{align*}
1 & \leftarrow 2 \rightarrow \cdots \rightarrow n-1 \rightarrow n \\
1 & \rightarrow 2 \rightarrow \cdots \rightarrow n-1 \rightarrow n
\end{align*}
$$

The latter we have seen in the context of persistent homology, the former is integral in the construction of levelset zigzag persistent homology.

**Example 12.1.** Let $K$ and $K'$ be two simplicial complexes indexed over the same index set and let $K \cap K'$ denote the simplex-wise intersection. This defines a zigzag of simplicial complex $K \leftarrow K \cap K' \rightarrow K'$, and a corresponding zigzag of vector spaces:

$$
H_i(K) \leftarrow H_i(K \cap K') \rightarrow H_i(K').
$$

The simplicial complexes in Fig. 38 give such an example. By choosing a basis

$$
\{\{1,2\} + \{1,3\} + \{2,3\} + \{2,4\} + \{3,4\}\}
$$

for $H_1(K \cap K')$, we get

$$
\mathbb{Z}_2 \xleftarrow{[1,0]} \mathbb{Z}_2 \oplus \mathbb{Z}_2 \xrightarrow{[0,1]} \mathbb{Z}_2 = \left( \mathbb{Z}_2 \xleftarrow{1} \mathbb{Z}_2 \rightarrow 0 \right) \oplus \left( 0 \xleftarrow{1} \mathbb{Z}_2 \rightarrow \mathbb{Z}_2 \right).
$$

The previous example illustrates the important fact that zigzag modules too have a well-defined barcodes. For a zigzag on $n$ vertices, let $[a,b]$ denote the restriction of the poset to all vertices $i \in [a,b]$. The associated **interval module** $I^{[a,b]}$ is defined by

$$
I^{[a,b]}_i = \begin{cases} 
k & \text{if } i \in [a,b] \\
0 & \text{otherwise}
\end{cases}
$$
together with the identify morphism id: \( I_1^{[a,b]} \rightarrow I_1^{[a,b]} \) whenever \( i, j \in [a,b] \). E.g. if \( n = 4 \), and the arrows alternate, then \( I_2^{[2,3]} \) is the persistence module

\[
0 \rightarrow k \leftarrow k \rightarrow 0.
\]

We state the following result without proof. A proof can be obtained by a variation of the proof of Theorem 6.16.

**Theorem 12.2.** Let \( V \) be a zigzag module on \( n \) vertices such that \( \dim V_p < \infty \) for all \( p \in [n] \). Then

\[
V \cong \bigoplus_{[a,b] \in B(V)} I^{[a,b]}
\]

where \( B(V) \) is a multiset of intervals in \([n]\) called the barcode of \( V \).

### 12.1.1 Levelset Zigzag Persistent Homology

The persistent homology of a continuous function \( f : X \rightarrow \mathbb{R} \) studies the evolution of the homology of the sublevel sets. An alternative approach would be to study how the homology changes across the levelsets \( f^{-1}(t) \) as the parameter \( t \) sweeps over the real line. Consider for instance the topological space in Fig. 39 with the real function being projection onto the horizontal axis. We notice two properties of the function:

- \( f^{-1}(a_i, a_{i+1}) \) is homeomorphic to \( f^{-1}(s_i, s_{i+1}) \times (a_i, a_{i+1}) \) with \( f \) being the projection onto the second component,
- \( f^{-1}(s_i, s_{i+1}) \) deformation detracts onto \( f^{-1}(a_{i+1}) \).

In particular, we can study the evolution of the homology of the levelsets by applying homology to the following zigzag of topological spaces:

\[
\begin{align*}
&f^{-1}(-\infty, s_1) & f^{-1}(s_1, s_2) & f^{-1}(s_2, s_3) & f^{-1}(s_3, s_4) & f^{-1}(s_4, \infty) \\
f^{-1}(s_1) & f^{-1}(s_2) & f^{-1}(s_3) & f^{-1}(s_4) & f^{-1}(s_4)
\end{align*}
\]

Proceeding in \( H_0 \) with the “obvious basis” we get the following zigzag of vector spaces

\[
\begin{align*}
&\mathbb{Z}_2 \oplus \mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 \oplus \mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 \oplus \mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 \oplus \mathbb{Z}_2 \\
&\mathbb{Z}_2 & 1 & \mathbb{Z}_2 & 1 & \mathbb{Z}_2 & 1 & \mathbb{Z}_2
\end{align*}
\]

For each copy of \( \mathbb{Z}_2 \oplus \mathbb{Z}_2 \) above, we replace current basis \( \{v_1, v_2\} \) with the basis \( \{v_1, v_1 + v_2\} \). The zigzag then diagonalizes as

\[
\begin{align*}
&\mathbb{Z}_2 \oplus \mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 \oplus \mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 \oplus \mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 \oplus \mathbb{Z}_2 \\
&\mathbb{Z}_2 & 1 & \mathbb{Z}_2 & 1 & \mathbb{Z}_2 & 1 & \mathbb{Z}_2
\end{align*}
\]

which is the direct sum of the following three zigzags:

\[
\begin{align*}
&\mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 & \mathbb{Z}_2 \\
&1 & 1 & 1 & 1 & 1 & 1 & 1
\end{align*}
\]
The corresponding bars will all be given in terms of the critical values $a_i$: For the first zigzag the feature is generated over $f^{-1}(-\infty, s_1] \simeq f^{-1}(a_1)$ and it lives up to and including $f^{-1}[s_4, \infty) \simeq f^{-1}(a_5)$. Therefore we identify this feature with the interval $[a_1, a_5]$; this represents the global connected component of the topological space. For the second zigzag we have an interval module generated over $f^{-1}(-\infty, s_1] \simeq f^{-1}(a_1)$ which vanishes as it enters $f^{-1}[s_1, s_2] \simeq f^{-1}(a_2)$. This is identified with the half-open interval $[a_1, a_2)$ and represents one of the two "legs" living over the interval $[a_1, a_2)$. Lastly we have an interval which lives over $f^{-1}(s_3)$ and vanishes as it enters $f^{-1}[s_2, s_3] \simeq f^{-1}(a_3)$ and $f^{-1}[s_3, s_4] \simeq f^{-1}(a_4)$. Hence, this yields an interval $(a_3, a_4)$ and its geometrical interpretation is clear from the figure. Denoting the resulting barcode by $ZZ_0(f)$ we have shown that

$$ZZ_0(f) = \{[a_1, a_5], [a_1, a_2), (a_3, a_4)\}.$$ 

And for $H_1$ (exercise)

$$ZZ_1(f) = \{[a_1, a_5], [a_1, a_2], [a_3, a_4]\}.$$ 

What about sublevel persistent homology? The intervals in dimension 0 and dimension 1 are, respectively

$$\{[a_1, \infty), [a_1, a_2)\} \quad \text{and} \quad \{[a_1, \infty), [a_1, a_5), [a_3, \infty), [a_4, \infty)\}.$$ 

These two barcodes seem closely related. We now state a theorem which makes this connection precise.

**Theorem 12.3 (See [4]).** The sublevel set barcode in dimension $n$ is given as may the following union:

$$\{[a_i, a_j) : [a_i, a_j) \in ZZ_n(f)\} \cup \{[a_i, \infty) : [a_i, a_j] \in ZZ_n(f)\} \cup \{[a_j, \infty) : (a_i, a_j) \in ZZ_{n-1}(f)\}.$$ 

Hence, levelset zigzag persistent homology contains strictly more information. Assuming functions of *morse type* - i.e. that the function has a finite set of critical value as in the previous example - one also has the following stability theorem.

**Theorem 12.4.** Let $f, g: X \to \mathbb{R}$, and assume that $\dim H_i(f^{-1}(t)), H_i(g^{-1}(t)) < \infty$ for all $t \in \mathbb{R}$. Then

$$d_B(ZZ_n(f), ZZ_n(g)) \leq ||f - g||_\infty.$$ 

Much more can be said about levelset zigzag persistent homology, but we will not pursue this.

**12.1.2 Reeb Graphs**

For Reeb Graphs the intervals in the levelset zigzag barcode have a clear geometric interpretation. We illustrate this with an example.
Example 12.5. Consider the Reeb graph shown in Fig. 10. It has 7 critical points valued \( a_1 < a_2 < a_3 < a_4 \). Choosing values \( s_i \) such that \( a_i < s_i < a_{i+1} \) yields the following zigzag of topological spaces:

\[
\begin{align*}
&f^{-1}(\infty, s_1) \quad f^{-1}(s_1, s_2) \quad f^{-1}(s_2, s_3) \quad f^{-1}(s_3, \infty) \\
\end{align*}
\]

Applying \( H_0 \) then produces the following zigzag of vector spaces:

\[
\begin{align*}
&Z_2 \oplus Z_2 \\
&\quad \xrightarrow{I_2} \quad Z_2 \\
&\quad \xrightarrow{[1,1]} \quad Z_2 \oplus Z_2 \\
&\quad \xrightarrow{\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \quad Z_2 \oplus Z_2 \\
&\quad \xrightarrow{I_2} \quad Z_2 \oplus Z_2 \\
&\quad \xrightarrow{I_2} \quad Z_2 \oplus Z_2 \\
\end{align*}
\]

The associated barcode is: \( ZZ_0(f) = \{[a_1, a_4], [a_1, a_2], (a_2, a_4), (a_2, a_3)\} \) with the interpretation

- \([a_1, a_4]\): the global connected component.
- \([a_1, a_2]\): the downward facing branch.
- \((a_2, a_4)\): the upward facing branch.
- \((a_2, a_3)\): the loop.

The interpretation of different types of intervals as particular graph theoretic features holds for Reeb graphs in general. While different (=non-isomorphic) Reeb graphs may exhibit the same barcode, the associated bottleneck distance still provides an efficiently computable similarity measure. This contrasts the more discriminative distances - such as an interleaving distance for Reeb graphs - which are generally NP-hard to compute.

### 12.2 Multiparameter Persistent Homology

Persistent homology is stable with respect to perturbation, but at the same time highly sensitive to noise and outliers. Take for instance the noisy sample of a circle depicted in Fig. 11: its associated barcode in dimension 1 will exhibit no significant features. One could attempt to rectify that by restricting to data points with a local density estimate greater than a predefined threshold \( \delta \). However, the resulting barcode would be highly sensitive to the choice of \( \delta \): chosen too large and there won’t be enough points to cover a circle, chosen too small and the data will be too noisy.
Another approach would be to consider all possible density thresholds at once and construct a $P$-module where $P$ is a grid. That is the idea behind **multiparameter persistent homology**.

Define $\mathbb{N} \times \mathbb{N}$ to be the partially ordered set consisting of pairs $(i,j) \in \mathbb{N} \times \mathbb{N}$ and $(i,j) \leq (i',j')$ if and only if $i \leq i'$ and $j \leq j'$. E.g. $[3] \times [2]$ is the following grid:

$$\begin{array}{ccc}
\bullet & \rightarrow & \bullet \\
\bullet & \rightarrow & \bullet \end{array}$$

Let $(Q,d)$ be a finite metric space and let $f: Q \rightarrow \mathbb{R}$ be any function. For a collection of function thresholds $\delta_1 > \ldots > \delta_m$ and distance thresholds $\epsilon_1 < \epsilon_2 < \ldots < \epsilon_n$ we get an $[n] \times [m]$-module $M$ by defining

$$M_{i,j} = H_i(\text{VR}_{\epsilon_i}(f^{-1}[\delta_j, \infty)))$$

and where the map $M_{i,j} \rightarrow M_{i',j'}$ is induced in homology by the inclusion

$$\text{VR}_{\epsilon_i}(f^{-1}[\delta_j, \infty)) \hookrightarrow \text{VR}_{\epsilon_{i'}}(f^{-1}[\delta_{j'}, \infty)).$$

**Example 12.6.** Let $Q = \{q_1, q_2, q_3\}$ with pairwise distances $d(q_2, q_3) < d(q_1, q_3) = d(q_1, q_2)$ and $f(q_1) = f(q_2) > f(q_3)$, as depicted in Fig. 42. The resulting $[3] \times [2]$-module in 0-dimension homology is:

$$\begin{align*}
\text{span}\{q_1, q_2, q_3\} & \xrightarrow{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix}} \text{span}\{q_1, q_2\} \xrightarrow{[1,1]} \text{span}\{q_1\} \\
\text{span}\{q_1, q_2\} & \xrightarrow{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \text{span}\{q_1, q_2\} \xrightarrow{[1,1]} \text{span}\{q_1\}
\end{align*}$$

Figure 40: A Reeb graph.
By an appropriate change of basis this diagram transforms to:

\[
\begin{align*}
\text{span}\{q_1, q_1 + q_2, q_2 + q_3\} & \xrightarrow{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \text{span}\{q_1, q_1 + q_2\} & \xrightarrow{[1,0]} \text{span}\{q_1\} \\
\text{span}\{q_1, q_1 + q_2\} & \xrightarrow{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}} \text{span}\{q_1, q_1 + q_2\} & \xrightarrow{[1,0]} \text{span}\{q_1\}
\end{align*}
\]

which is isomorphic the direct sum of the following three modules:

\[
\begin{align*}
Z_2 & \xrightarrow{1} Z_2 & \xrightarrow{1} Z_2 & \xrightarrow{1} Z_2 & \xrightarrow{1} Z_2 & \rightarrow 0 \\
& \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\
Z_2 & \xrightarrow{1} Z_2 & \xrightarrow{1} Z_2 & \xrightarrow{1} Z_2 & \xrightarrow{1} Z_2 & \rightarrow 0 \\
& \uparrow & \uparrow & \uparrow & \uparrow & \uparrow \\
& & & & & & \rightarrow 0 \\
& & & & & & \rightarrow 0 \\
& & & & & & \rightarrow 0
\end{align*}
\]
With the previous example fresh in mind one may be tempted to conjecture that one can decompose multiparameter persistence modules into simple components as in Theorem 12.2. That turns out to be very far from the truth.

**Example 12.7.** By applying $H_1$ to the diagram depicted in Fig. 43 one arrives at the following commutative diagram of vector spaces

\[
\begin{array}{c}
\mathbb{Z}_2 \\ [0,1]
\end{array} \rightarrow \begin{array}{c}
0 \\ [1,1]
\end{array} \rightarrow \begin{array}{c}
0
\end{array} \\
\begin{array}{c}
\mathbb{Z}_2[1,1]
\end{array} \rightarrow \begin{array}{c}
\mathbb{Z}_2 \\ [1,1]
\end{array} \rightarrow \begin{array}{c}
0
\end{array} \\
\begin{array}{c}
\mathbb{Z}_2[1,1]
\end{array} \rightarrow \begin{array}{c}
\mathbb{Z}_2 \\ [1,1]
\end{array} \rightarrow \begin{array}{c}
\mathbb{Z}_2[1,0]
\end{array} \rightarrow \begin{array}{c}
\mathbb{Z}_2
\end{array}
\]

One can show that this diagram is indecomposable: no matter what change of basis you perform, the diagram will not split up into simpler components.

The previous example illustrates an algebraic fact that makes multiparameter persistent homology significantly more involved than the one-parameter case. While it follows from 6.12 that any $[n] \times [m]$-module decomposes into a direct sum of indecomposable modules, one does not have a reasonable way of parametrizing all the indecomposable representations which could potentially arise. Granted, for a specific module $M$, one can decompose $M$ into its indecomposable summands (up to isomorphism). How to use such indecomposable modules in data analysis is not clear and is an active topic of research.

### 12.2.1 RIVET

Rather than seeking to understand the full complexity of multiparameter persistence modules, one can study them through the lens of restrictions to linear subposets. Specifically, assuming that $M$ is an $\mathbb{R}^2$-indexed persistence module, we can restrict $M$ to an arbitrary straight line $L$ in the plane with non-negative slope. The restricted module $M|_L$ has a well-defined barcode by means of Theorem 6.18, assuming that the vector space dimensions are finite. As $L$ ranges over all valid lines we can gain valuable insight the underlying 2-parameter persistence module $M$.

RIVET (the Rank Invariant Visualization and Exploration Tool) allows for the visualization and analysis of two-parameter persistence modules, and arguably its main feature is the fast computation of $M|_L$ for all lines $L$. The algorithm is rather involved and requires a discussion of commutative algebra beyond the scope of this course.

---

Example 12.8. Let \( Q \) be the points in Fig. 41 and let \( f : Q \to \mathbb{R} \) be the function which for each point \( q \) counts the number of data points within unit distance. Associated to this data RIVET considers the \( \mathbb{R}^2 \)-indexed persistence module \( M \) given by

\[
M(s, t) = \text{VR}_t(f^{-1}[s, \infty)).
\]

Be aware that the density filtration takes place in the horizontal direction, and that the scale parameter in RIVET - as in Ripser - is twice the one in the definition of the Vietoris–Rips complex used in these notes. Consider the subset of \( Q \) admitting a local density estimate at least 18, as seen in Fig. 44c. Its associated barcode in the Rips filtration is displayed along the thick blue line in Fig. 45c. We see that there is a short bar generated at a fairly large scale corresponding to the scale at which the right-end of the point cloud connects. For a much cleaner barcode we consider all the points with local density estimate at least 10, as shown in Fig. 44b together with its barcode in Fig. 45b. It is often more fruitful to consider the barcode along a slanted line; see Fig. 45a. RIVET updates the barcode in real-time as the user modifies the slope and intercept.

12.2.2 Distances

With the failure of a barcode-like decomposition for multiparameter persistence module the question of how to compare modules arises. We now give an informal introduction to two
well-studied distances in the context of multiparameter persistent homology.

The matching distance  Let $L$ be a line with positive slope in $\mathbb{R}^2$ and let $M$ and $N$ be $\mathbb{R}^2$-indexed persistence modules satisfying $\dim M_{(s,t)} < \infty$ and $\dim N_{(s,t)} < \infty$ for all $(s,t)$. Assuming that $L$ is described by $ax + b$ for some $a > 0$, we define the distance between $M$ and $N$ along $L$ to be the following weighted bottleneck distance:

$$F_{M,N}(L) = w(L)d_B(M|_L, N|_L)$$

where $w(L)$ is given by

$$w(L) := \begin{cases} 
\frac{1}{\sqrt{1+a^2}} & a \geq 1 \\
\frac{1}{\sqrt{1+a^2}} & a \leq 1 
\end{cases}$$

We define the matching distance by:

$$d_{\text{match}}(M, N) = \sup_L F_{M,N}(L).$$

Remark 12.9. The introduction of a weight in the supremum is order to make sure that the distance is ”stable” in a precise way. We omit the details.

While the matching distance can be approximated efficiently, a major drawback is that it is a rather insensitive distance measure.

The interleaving distance  We follow the definition of an interleaving in [9, 10] Define the $\epsilon$-shift of an $\mathbb{R}^2$-indexed persistence module $M$ to be the persistence module $M^\epsilon$ defined by $M^\epsilon_{(s,t)} = M_{(s,t)} + (\epsilon, \epsilon)$ and $M^\epsilon_{((s,t) \leq (s',t'))} = M_{((s,t) + (\epsilon, \epsilon) \leq (s',t')}$. If $f : M \to N$ is a morphism, then we get an $\epsilon$-shifted morphism $f^\epsilon : M^\epsilon \to N^\epsilon$ defined by $f^\epsilon_{(s,t)} = f_{(s,t) + (\epsilon, \epsilon)}$. Let $\eta^\epsilon_M : M \to M^\epsilon$ be the morphism whose restriction to each $M_{(s,t)}$ is the internal morphism $M_{((s,t) \leq (s',t') + (\epsilon, \epsilon))}$.

Definition 12.10. Given $\epsilon \in [0, \infty)$, an $\epsilon$-interleaving between $M$ and $N$ is a pair of morphisms $\psi : M \to N^\epsilon$ and $\varphi : N \to M^\epsilon$ such that $\varphi^\epsilon \circ \psi = \eta^\epsilon_M$ and $\psi^\epsilon \circ \varphi = \eta^\epsilon_N$. We say that $M$ and $N$ are $\epsilon$-interleaved.

This definition generalizes to more parameters in a straightforward way.

Michael Lesnick[9] showed that the interleaving distance is the most discriminative and stable distance measure on multiparameter persistence modules[5]. Unfortunately, it turns out that its computation is NP-hard.

Theorem 12.11 (Bjerkevik-Botnan-Kerber[3]). It is NP-hard to approximate the interleaving distance within a factor of $3$.

The proof proceeds by a reduction from the following problem which is subsequently shown to be NP-complete: let $A$ and $B$ be $n \times n$-matrices and assume that a pre-assigned subset of the indices of $A$ and $B$ are set 0. Is it possible to fill in the remaining indices such that the resulting product $AB$ is the identify matrix $I_n$?

Example 12.12. Let

$$A = \begin{bmatrix} * & * & * \\ * & 0 & * \\ * & * & 0 \end{bmatrix} \quad B = \begin{bmatrix} * & * & * \\ * & 0 & 0 \\ * & 0 & 0 \end{bmatrix}.$$
In this case the decision problem has a positive answer:

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 0 & 1 \\
1 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
-1 & 1 & 1 \\
1 & -1 & 0 \\
1 & 0 & -1
\end{bmatrix}
= I_3.
\]
13 Week 14: $d_I = d_B$

In this section we will prove Theorem 9.13 by showing that $d_I \leq d_B$ and $d_B \leq d_I$. The former inequality first appeared in [9] and is rather straightforward. For the latter we will be following the work of Bjerkevik [2]. Other proofs do exist in the literature but we will focus on this particular approach because 1) it is combinatorial in nature, 2) it generalizes to multiparameter modules in a way that other approaches do not. We shall however work under the additional assumption of a finite number of intervals. This is just for notational simplicity as essentially the same proof works in the more general setting of Theorem 9.13.

In the second part of the proof we shall make two additional assumptions for the sake of brevity:

1. All the intervals have finite support.
2. Every interval is of the form $[a, b)$.

Under the assumptions of a finite number of intervals, the first of these can be made without loss of generality. Indeed, one can replace infinite intervals with intervals of the form $[-u, b)$, $[u, \infty)$ and $[-u, u)$ for $u$ sufficiently large. The latter assumption can also be made without loss of generality as the following lemma shows.

**Lemma 13.1.** Let $M \cong \oplus_{(a,b) \in B(M)} I^{(a,b)}$ and let $\hat{M} = \oplus_{(a,b) \in B(M)} I^{(a,b)}$. Then

$$d_B(B(M), B(\hat{M})) = 0, \quad d_I(M, \hat{M}) = 0.$$  

**Proof.** Exercise.

Hence, from the triangle inequality it follows that $d_B(B(M), B(N)) = d_B(B(\hat{M}), B(\hat{N}))$ and $d_I(M, N) = d_I(\hat{M}, \hat{N})$.

13.1 $d_I \leq d_B$

Recall that for an interval $J \subseteq \mathbb{R}$ we have the interval module $I^J$ defined by

$$I^J_t = \begin{cases} k & \text{if } t \in J \\ 0 & \text{otherwise.} \end{cases}$$

and with $I^J(t \leq t')$ the identity morphism whenever $t, t' \in J$.

**Lemma 13.2.** Let $J$ and $K$ be intervals in $\mathbb{R}$. Then,

$$d_B(\{J\}, \{K\}) = d_I(I^J, I^K),$$

and

$$d_B(\{J\}, \emptyset) = d_I(I^J, 0).$$

**Proof.** Exercise.

**Lemma 13.3.** If $M$ and $N$ are $\epsilon$-interleaved, and $M'$ and $N'$ are $\epsilon$-interleaved, then $M \oplus M'$ and $N \oplus N'$ are $\epsilon$-interleaved.
Proof. Let \( f : M \to N^\epsilon \) and \( g : N \to M^\epsilon \), and \( f' : M' \to (N')^\epsilon \) and \( g' : N' \to (M')^\epsilon \) constitute \( \epsilon \)-interleavings. We thus get morphisms

\[
\begin{bmatrix}
    f & 0 \\
    0 & f'
\end{bmatrix}: M \oplus M' \to N^\epsilon \oplus (N')^\epsilon
\]

and

\[
\begin{bmatrix}
    g & 0 \\
    0 & g'
\end{bmatrix}: N \oplus N' \to M^\epsilon \oplus (M')^\epsilon.
\]

It remains to check that they constitute an \( \epsilon \)-interleaving. That is straightforward:

\[
\begin{bmatrix}
    g & 0 \\
    0 & (g')^\epsilon
\end{bmatrix} \cdot \begin{bmatrix}
    f & 0 \\
    0 & f'
\end{bmatrix} = \begin{bmatrix}
    \eta_{M}^{2\epsilon} & 0 \\
    0 & \eta_{M'}^{2\epsilon}
\end{bmatrix} = \eta_{M \oplus M'}^{2\epsilon}.
\]

Likewise,

\[
\begin{bmatrix}
    f & 0 \\
    0 & (f')^\epsilon
\end{bmatrix} \cdot \begin{bmatrix}
    g & 0 \\
    0 & g'
\end{bmatrix} = \eta_{N \oplus N'}^{2\epsilon}.
\]

Combining the two previous lemmas we arrive at our result:

**Theorem 13.4.** Let \( B(M) \) and \( B(N) \) be a finite multisets of intervals. Then \( d_I(M, N) \leq d_B(B(M), B(N)) \).

**Proof.** Let \( \chi \) be a matching between \( d_B(B(M), B(N)) \) with cost \( c(\chi) = \epsilon \). If \( J \in B(M) \) is matched to \( K \in B(N) \), then \( d_B([J], [K]) \leq C(J, K) = \epsilon \), and by Lemma 13.2 there is an \( (\epsilon + \delta) \)-interleaving between \( I^J \) and \( I^K \) for an arbitrary small \( \delta > 0 \). Likewise, if \( J \in B(M) \cup B(N) \) is unmatched, then \( I^J \) is \( (\epsilon + \delta) \)-interleaved with the 0-module for an arbitrary small \( \delta > 0 \). Hence, by invoking Lemma 13.3 we conclude that

\[
M \cong \left( \bigoplus_{J \in B(M) \text{ matched}} I^J \right) \bigoplus \left( \bigoplus_{J' \in B(M) \text{ unmatched}} I^{J'} \right) \bigoplus 0
\]

is \( (\epsilon + \delta) \)-interleaved with

\[
N \cong \left( \bigoplus_{K \in B(N) \text{ matched}} I^K \right) \bigoplus 0 \bigoplus \left( \bigoplus_{K' \in B(N) \text{ unmatched}} I^{J'} \right)
\]

Since \( \delta > 0 \) is arbitrary small, it follows that \( d_I(M, N) \leq \epsilon \). \( \Box \)

The more general case where \( B(M) \) and \( B(N) \) contain an infinite number of intervals follows by showing that Lemma 13.3 generalizes to arbitrary direct sums. We leave that as an exercise.

**13.2 \( d_B \leq d_I \)**

In the following, every interval is assumed to be of the form \([a, b)\) for \( b < \infty \).

**Lemma 13.5.** If \( I_1 = [a_1, b_1) \) and \( I_2 = [a_2, b_2) \) are intervals and \( f : I^{I_1} \to I^{I_2} \) is non-zero, then \( a_2 \leq a_1 \) and \( b_2 \leq b_1 \).
Proof. Assume \( a_2 > a_1 \). The following diagram must necessarily commute for any \( t \in [a_1, b_1) \):

\[
\begin{array}{c}
I_{a_1}^{I_1} = k \\
\downarrow f_{a_1} \\
I_{a_1}^{I_2} = 0 \\
\downarrow f_t \\
I_t^{I_2} = k
\end{array}
\]

Since the morphism \( f_{a_1} = 0 \), it follows that \( f_t = 0 \). Hence \( f \) is the zero-morphism. A similar argument shows that \( b_2 \leq b_1 \). \( \square \)

**Lemma 13.6.** Any morphism \( f : I^{[a_1, b_1]} \to I^{[a_2, b_2]} \) is determined by \( f_{a_1} \in k \).

**Proof.** The morphism is trivial outside of \([a_1, b_2]\) by the previous lemma. For any \( t \in [a_1, b_2) \) we have the following commutative diagram

\[
\begin{array}{c}
I_{a_1}^{[a_1, b_1]} = k \\
\downarrow f_{a_1} \\
I_{a_1}^{[a_1, b_2]} = k \\
\downarrow f_t \\
I_t^{[a_1, b_2]} = k
\end{array}
\]

Hence \( f_t = f_{a_1} \). \( \square \)

For an interval \( I = [a, b] \) define \( \alpha(I) = a + b \). This defines a preorder on intervals: \( I \leq J \) if \( \alpha(I) \leq \alpha(J) \).

**Lemma 13.7.** Let \( I_1, I_2 \) and \( I_3 \) be intervals, \( \alpha(I_1) \leq \alpha(I_3) \), and \( 0 \neq f : I^{I_1} \to (I^{I_2})^\epsilon \) and \( 0 \neq g : I^{I_2} \to (I^{I_3})^\epsilon \). Then \( I^{I_2} \) is \( \epsilon \)-interleaved with \( I^{I_1} \) or \( I^{I_3} \).

**Proof.** Write \( I_1 = [a_1, b_1] \), \( I_2 = [a_2, b_2] \) and \( I_3 = [a_3, b_3] \). From Lemma 13.5

\[
\begin{align*}
\alpha(I_1) &= a_1 + b_1 \\
\alpha(I_2) &= a_2 + b_2 \\
\alpha(I_3) &= a_3 + b_3
\end{align*}
\]

Assume that \( I^{I_1} \) and \( I^{I_2} \) are not \( \epsilon \)-interleaved. The modules \( I^{[a_1, b_1]} \) and \( I^{[a_2, b_2]} \) are \( \epsilon \)-interleaved if \( |a_1 - a_2| \leq \epsilon \) and \( |b_1 - b_2| \leq \epsilon \), and thus we may assume that \( a_2 < a_1 - \epsilon \) or \( b_2 < b_1 - \epsilon \). Assuming the former:

\[
\alpha(I_1) = a_1 + b_1 > a_2 + \epsilon + b_1 \geq a_2 + b_2 = \alpha(I_2),
\]

and assuming the latter:

\[
\alpha(I_1) = a_1 + b_1 > a_1 + b_2 + \epsilon \geq a_2 + b_2 = \alpha(I_2).
\]

In either case, \( \alpha(I_1) > \alpha(I_2) \). Assuming that \( I^{I_2} \) and \( I^{I_3} \) are not \( \epsilon \)-interleaved and by substituting \( I_2 \) for \( I_1 \), and \( I_3 \) for \( I_2 \) in the above argument, we get \( \alpha(I_2) > \alpha(I_3) \). In conclusion:

\[
\alpha(I_1) > \alpha(I_2) > \alpha(I_3),
\]

a contradiction. \( \square \)

**Lemma 13.8.** Let \( I_1 = [a_1, b_1] \), \( I_2 = [a_2, b_2] \) and \( I_3 = [a_3, b_3] \) be such that \( b_1 - a_1 > 2\epsilon \), \( b_3 - a_3 > 2\epsilon \) and \( \alpha(I_1) \leq \alpha(I_3) \). If \( 0 \neq f : I^{I_1} \to (I^{I_2})^\epsilon \) and \( 0 \neq g : I^{I_2} \to (I^{I_3})^\epsilon \). Then \( g^\epsilon \circ f \neq 0 \).
Using the aforementioned notation we see that we can rewrite this as:

\[ \pi \]

Furthermore, by definition of a direct sum, \( 0 = (\pi^M)^{2\epsilon} \circ \eta^M \circ i^M \) whenever \( i \neq i' \). In turn,

\[ 0 = \sum_{J \in B(N)} g_{J,I}^i \circ f_{I,J}. \]
Proof of Lemma 13.9. Order the elements of $A = \{I_1, \ldots, I_r\}$ such that $\alpha(I_i) \leq \alpha(I_i')$ for $i \leq i'$, and let $\mu(A) = \{J_1, \ldots, J_s\}$. By combining Lemma 13.7 with Eq. (10), we arrive at

$$\eta_{ij}^2 = \sum_{J \in B(N)} g'_{I,J} \circ f_{I,J} = \sum_{J \in \mu(A)} g'_{I,J} \circ f_{I,J}$$

This is true because $f_{I,J} = 0$ or $g_{I,J} = 0$ for every $J \not\in \mu(I)$. Likewise, assuming $i < i'$, Lemma 13.7 together with Eq. (11) yield

$$0 = \sum_{J \in B(N)} g'_{I,J'} \circ f_{I,J} = \sum_{J \in \mu(A)} g'_{I,J'} \circ f_{I,J}.$$

For $I = [a,b] \in A$, Lemma 13.6 states that every morphism $h: I^f \to I^K$ is determined by the constant $h_a$. Let $w(h) \in k$ denote this constant. Clearly, $w(f_{I,J}) = w(f_{I,J'})$. If $i \leq i'$, then Lemma 13.8 asserts that

$$w(g'_{I,J'} \circ f_{I,J}) = w(g'_{I,J'}) \cdot w(f_{I,J}) = w(g_{I,J'}) \cdot w(f_{I,J}).$$

Hence,

$$1 = w(\eta_{ij}^2) = \sum_{J \in \mu(I)} w(g_{I,J}) \cdot w(f_{I,J})$$

and

$$0 = \sum_{J \in \mu(I)} w(g_{I,J}) \cdot w(f_{I,J})$$

for $i < i'$. Putting this in matrix form:

$$\begin{bmatrix} w(g_{I_1,J_1}) & \cdots & w(g_{I_1,J_s}) \\ \vdots & \ddots & \vdots \\ w(g_{I_r,J_1}) & \cdots & w(g_{I_r,J_s}) \end{bmatrix} \cdot \begin{bmatrix} w(f_{I_1,J_1}) & \cdots & w(f_{I_1,J_s}) \\ \vdots & \ddots & \vdots \\ w(f_{I_r,J_1}) & \cdots & w(f_{I_r,J_s}) \end{bmatrix} = \begin{bmatrix} 1 & ? & \cdots & ? \\ 0 & 1 & ? & \cdots & ? \\ 0 & 0 & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

where we the products on the upper diagonal are unknown. The right hand side has rank $r = \|A\|$, whereas the the rank of the left hand side is bounded by $s = \|\mu(A)\|$. We conclude that $\|A\| \leq \|\mu(A)\|$.

13.2.1 Hall’s Theorem

Consider a hypothetical course where students have to present a particular paper, but no two students can work on the same paper. Each student naturally has a preference, and the problem is whether or not one can assign papers such that all students are satisfied. Let $P$ denote the set of papers and let $A_j \subseteq P$ denote the set of paper student $j$ would like to study.

Example 13.10. Let $P = \{1, 2, 3, 4, 5\}$ and $A_1 = \{1, 2, 3\}$, $A_2 = \{1, 4, 5\}$ and $A_3 = \{3, 5\}$. In this case Student 1 can work on Paper 1, Student 2 can work on Paper 4 and Student 3 can work on Paper 5.

Example 13.11. Let $P = \{1, 2, 3, 4, 5\}$ and $A_1 = \{2, 3, 4, 5\}$, $A_2 = \{4, 5\}$, $A_3 = \{4\}$ and $A_5 = \{5\}$. In this case no assignment is possible.
Let $G$ be a bipartite graph on sets $X$ and $Y$, and for a subset of vertices $X' \subseteq X$,

$$N_G(X') = \{ w \in G : \text{ there is an edge between } w \text{ and } v \in X' \}.$$ 

**Theorem 13.12** (Hall's Theorem). *The following two are equivalent:

1. There exists an injection $\chi : X \mapsto Y$ such that $\chi(v) = w$ only if there is an edge between $v$ and $w$.
2. $|X'| \leq |N_G(X')|$ for all $X' \subseteq X$.

**Example 13.13.** Returning to Example 13.11 we see let $X = \{S_1, S_2, S_3, S_4, S_5\}$ and $Y = \{1, 2, 3, 4, 5\}$, and we connect $S_i \in X$ with $j \in Y$ if $j \in A_i$. There is no way of assigning papers as $3 = |\{S_3, S_4, S_5\}| > |N_G(\{S_3, S_4, S_5\})| = |\{4, 5\}| = 2$.

### 13.2.2 Wrapping Up

Let $B(M)_{2\epsilon} = \{(a, b) \in B(M) : b - a > 2\epsilon\}$ and $B(N)_{2\epsilon} = \{(a, b) \in B(N) : b - a > 2\epsilon\}$. Recalling the definition of the cost of a matching from Eq. (5), we see that for intervals $I$ and $J$

$$d_B(\{I\}, \{J\}) = \min\{\max\{c(I), c(J)\}, c(I, J)\}.$$ 

If $I \in B(M)_{2\epsilon}$ or $J \in B(M)_{2\epsilon}$ and $d_B(\{I\}, \{J\}) = d_I(I^I, I^J) \leq \epsilon$, it must thus be true that $c(I, J) \leq \epsilon$.

**Theorem 13.14.** Let $B(M)$ and $B(N)$ be finite multisets of finite intervals. Then,

$$d_B(B(M), B(N)) \leq d_I(M, N).$$

**Proof.** Assume that $M$ and $N$ are $\epsilon$-interleaved. Combining Lemma 13.9 with Theorem 13.12 we get injections $\chi_1 : B(M)_{2\epsilon} \mapsto B(N)$ and $\chi_2 : B(N)_{2\epsilon} \mapsto B(M)$ such that

$$\chi_1(I) = J \Rightarrow I^I \text{ and } I^J \text{ are } \epsilon\text{-interleaved}$$

$$\chi_2(J) = I \Rightarrow I^I \text{ and } I^J \text{ are } \epsilon\text{-interleaved.}$$

In either case, $c(I, J) \leq \epsilon$.

We now explain how these combine to yield a matching $\chi$ between $B(M)$ and $B(N)$ with cost at most $\epsilon$. Form a bipartite graph $G$ with vertex set $B(M) \cup B(N)$ and with a directed edge from $I \rightarrow J$ if $\chi_1(I) = J$ or $\chi_2(I) = J$. Note that there is a directed edge going out of any interval in $B(M)_{2\epsilon} \cup B(N)_{2\epsilon}$.

We define a matching for each connected component $C_i$ of $G$ as follows:

1. If $C_i$ forms a directed cycle, add $(I, \chi_1(I))$ to $\chi$ for all $I \in C_i \cap B(M)$. This matches every element of $C_i$, and $c(I, \chi_1(I)) \leq \epsilon$.

2. If $C_i$ does not form a directed cycle, then $C_i$ must contain an initial interval $J$ and a terminal interval $K$. If $J \in B(M)$, then add $(I, \chi_1(I))$ to $\chi$ for all $I \in C_i \cap B(M)$. Otherwise, add $(\chi_2(I), I)$ to $\chi$ for all $I \in C_i \cap B(N)$. This matches all intervals in $C_i$ with the possible exception of $K$. As $K$ is not the source of a directed edge, $K \in B(M) \setminus B(M)_{2\epsilon} \cup B(N) \setminus B(N)_{2\epsilon}$. Therefore $K$ can be left unmatched at a cost of $c(K) \leq \epsilon$.

The resulting matching $\chi$ between $B(M)$ and $B(N)$ is an $\epsilon$-matching. We conclude that

$$d_B(B(M), B(N)) \leq d_I(M, N).$$
References


